Single cell transcriptomics of the developing zebrafish lens and identification of putative controllers of lens development

Dylan Farnsworth[a], Mason Posner[b], Adam Miller[a]
[a]Institute of Neuroscience, University of Oregon, Eugene, OR, USA
[b]Department of Biology and Toxicology, Ashland University, Ashland, OH, USA

Corresponding author:
Mason Posner[b]
Ashland University
Department of Biology and Toxicology
401 College Avenue
Ashland, OH 44805
Email address: mposner@ashland.edu

Declarations of interest: none

Abstract

The vertebrate lens is a valuable model system for investigating the gene expression changes that coordinate tissue differentiation due to its inclusion of two spatially separated cell types, the outer epithelial cells and the deeper denucleated fiber cells that they support. Zebrafish are a useful model system for studying lens development given the organ's rapid development in the first several days of life in an accessible, transparent embryo. While we have strong foundational knowledge of the diverse lens crystallin proteins and the basic gene regulatory networks controlling lens development, no study has detailed gene expression in a vertebrate lens at single cell resolution. Here we report an atlas of lens gene expression in zebrafish embryos at single cell resolution through five days of development, identifying a number of novel regulators of lens development as potential targets for future functional studies. Our temporospatial expression data address open questions about the function of α-crystallins during lens development and provides the first detailed view of β- and γ-crystallin expression in and outside the lens. We describe subfunctionalization in transcription factor genes that occur as paralog pairs in the zebrafish. Finally, we examine the expression dynamics of cytoskeletal, RNA-binding, and transcription factors genes, identifying a number of novel patterns. Overall these data provide a foundation for identifying and characterizing lens developmental regulatory mechanisms and revealing targets for future functional studies with potential therapeutic impact.

Keywords: lens; development; zebrafish; scRNA-seq; crystallins; transcriptomics

1. Introduction

The vertebrate ocular lens is a valuable model system for studying development, physiology, disease and gene evolution. The presence of two terminal cell types, an outer single-layered epithelium and inner fiber cell mass, simplifies the examination of signaling processes leading to tissue differentiation (Grainger, 1992). The ease with which the lens can be imaged, both intact and histologically, makes it an excellent tool for identifying anatomical abnormalities and imaging cellular structure. Additionally, the co-option of numerous non-eye related genes during the early evolution of the vertebrate lens provides a unique opportunity to study the evolution of development and of proteins for diverse roles (Wistow and Piatigorsky, 1988). Most notably, small heat shock proteins became the α-crystallins of the lens, serving both as structural refractive elements while also inhibiting protein aggregation that would otherwise lead to lens cataract, the leading cause of human blindness worldwide.

Zebrafish have been an important species for studying eye development and function (Bibliowicz et al., 2011; Morris, 2011). The evolutionary conservation of gene expression and protein content between the zebrafish and mammalian lens facilitates the use of zebrafish as a model system for vision research (Greiling et al., 2009; Posner et al., 2008; Richardson et al., 2016). Some differences in eye development between fish and mammals, such as the mechanism of lens delamination from the surface ectoderm and the presence of duplicated gene paralogs, allows for comparative studies (Clemens et al., 2013; Greiling et al., 2010). Zebrafish are also an excellent model for mutational analysis due to their abundant and accessible embryos and availability of tools for gene editing (J. Liu et al., 2016). Gene targeting techniques have been used in zebrafish to interrogate the function of lens-associated crystallins (Mishra et al., 2018; Posner et al., 2013; Zou et al., 2015), aquaporins (Clemens et al., 2013; Froger et al., 2010; Vorontsova et al., 2018), transcription factors (Krall and Lachke, 2018; Shi et al., 2005; 2006), DNA methylating enzymes (Tittle et al., 2011), cytoskeletal proteins and most recently RNA-binding proteins (Shao et al., 2020).

Studies with mice have contributed to a strong foundational understanding of the molecular mechanisms underlying lens development. For example, mouse microarray gene expression data was used to develop the iSyTE database of genes with eye preferred expression, leading to the discovery of several novel lens-related transcription factors and the first known RNA-binding protein regulating lens development, TDRD7 (Agrawal et al., 2015; Audette et al., 2016; Kakrana et al., 2017; Lachke et al., 2011). RNA-Seq and proteomic analysis of mouse lenses microdissected to separate the outer, single layer epithelium from the inner elongated fiber cells provided detailed analyses of changes in gene expression through four developmental time points and identified multiple genes with increased expression in one of these two lens tissues (Zhao et al., 2018a; 2018b). In total, these studies validate an approach of using lens preferred gene expression as a tool for identifying potential novel developmental regulators that can then be tested by mutational analysis.

While these past studies in mouse, zebrafish and other vertebrate species have provided important insights into lens biology and development, current techniques that measure gene expression in individual cells open up new ways to visualize tissue preferred expression and dynamic changes in gene expression during tissue differentiation. No study to date has analyzed

lens development at single cell resolution. We recently published a single cell RNA-Seq (scRNA-Seq) atlas for zebrafish development, with over 44,000 cells from embryos at 1, 2 and 5 days post fertilization in 220 clusters annotated for specific tissue types (Farnsworth et al., 2019). Two of these clusters were identified as cells of the ocular lens. Here, we use this scRNA-Seq dataset to analyze changes in gene expression through these five days of zebrafish development. These embryonic stages span the time when the peripheral lens epithelium transitions to a simple cuboidal layer that both supports deeper fiber cells while also generating newly differentiating fiber cells at its posterior margin. We describe patterns of gene expression in the numerous and diverse crystallins, cytoskeletal and membrane proteins that contribute to lens transparency and the refractive power required for vision. We also compare the expression of known lens development regulatory genes between zebrafish and mammals and examine the impacts of duplicated zebrafish paralogs on regulatory networks. Our identification of genes with preferred expression in zebrafish lens epithelium and fiber cells provides a list of possible new targets for investigation of lens development, differentiation, and function.

2. Materials and methods

2.1. Fish husbandry

Fish were maintained by the University of Oregon Zebrafish facility using standard husbandry techniques (Westerfield, 2007) under an Institutional Animal Care and Use Committee (IACUC) approved protocol AUP-18-31. Zebrafish, *Danio rerio*, were bred and maintained at 28°C on a 14 hr on and 10 hr off light cycle. Animal care was provided by the University of Oregon fish facility staff, and veterinary care was provided by Dr. Kathy Snell, DVM. Zebrafish are not yet sexually determined at these stages of development (Wilson et al., 2014). Embryos were collected from natural matings, staged and pooled (n = 15 per replicate). Animals used in this study were: *Tg(olig2:GFP)vu12* for 1 dpf, n = 2; 2 dpf, n = 1; 5 dpf, n = 2 samples and *Tg(elavl3:GCaMP6s*) for 2 dpf, n = 1 sample.

2.2. Embryo dissociation

Collagenase P was prepared to a 100 mg/mL stock solution in HBSS. Chemical dissociation was performed using 0.25% Trypsin, Collagenase P (2 mg/mL), 1 mM EDTA (pH 8.0), and PBS for 15 min at 28C with gently pipetting every 5 min. Dissociation was quenched using 5% calf serum, 1 mM CaCl2, and PBS. Cells were washed and resuspended in chilled (4C), 1% calf serum, 0.8 mM CaCl2, 50 U/mL penicillin, 0.05 mg/mL streptomycin, and DMEM and passed through a 40 µM cell strainer (Falcon) and diluted into PBS + 0.04% BSA to reduce clumping. A final sample cell concentration of 2000 cells per microliter, as determined on a Biorad TC20 cell counter, was prepared in PBS +0.04% BSA for cDNA library preparation.

2.3. Single-cell cDNA library preparation

Sample preparation was performed by the University of Oregon Genomics and Cell Characterization core facility (https://gc3f.uoregon.edu/). Dissociated cells were run on a 10X Chromium platform using 10x v.2 chemistry targeting 10,000 cells. The resulting cDNA libraries

were amplified with 15 cycles of PCR and sequenced on either an Illumina HiSeq (5/6 samples) or an Illumina NextSeq (n = 1, 48 h dpf sample).

2.4. Computational analysis

The resulting sequencing data were analyzed using the 10X Cellranger pipeline, version 2.2.0 (Zheng et al., 2019) and the Seurat (Satija et al., 2015) software package for R, v3.4.4 (R Core Team, n.d.) using standard quality control, normalization, and analysis steps. Briefly, we aligned reads to the zebrafish genome, GRCz11_93, and counted expression of protein coding reads. The resulting matrices were read into Seurat where we performed PCA using 178 PCs. UMAP analysis was performed on the resulting dataset with 178 dimensions and a resolution of 18.0. For each gene, expression levels are normalized by the total expression, multiplied by a scale factor (10,000) and log-transformed. Differential gene expression analysis was performed using the FindMarkers function in Seurat using Wilcoxon rank sum test.

The lens cell clusters in the atlas were identified by mapping the expression of well-characterized lens crystallin and transcription factor genes. Fiber cell preferred β-crystallins and integral membrane proteins were used to identify that cell type while a known lens epithelial specific transcription factor, *foxe3*, was used to identify two epithelial cell clusters.

2.5. Analysis of gene expression shift in epithelial cells

We used gene ontology term (GO) analysis to compare the function of preferentially expressed genes between 1 and 2/5 dpf epithelial cells. The single cell atlas data were used to generate lists of genes with expression in our two identified epithelial cell clusters. Genes expressed in each cluster with adjusted p-values ≤ 0.05 were submitted to GOrilla to generate GO terms using the *Process* option (Eden et al., 2007; 2009).

3. Results

3.1. Well-characterized marker genes identified distinctly clustered lens epithelial and fiber cells

The 10X Genomics platform was used to identify and quantify mRNAs from 44,102 cells isolated from 1, 2 and 5 days post fertilization zebrafish embryos. Annotation of 220 cell clusters produced by UMAP non-linear dimensional reduction previously identified a lens cell region based on expression of lens specific crystallin protein genes (Farnsworth et al., 2019). Our first goal was to determine whether we could identify specific clusters for lens epithelial and fiber cells based on expression of known marker genes. For example, three genes known to be lens fiber cell enriched in mammals are the β-crystallin *cryba1a*, lactose binding *grifin*, and the aquaporin *mipa* (Hall and Mathias, 2014; Ogden et al., 1998; Wistow and Piatigorsky, 1988). These three genes showed shared expression in a specific set of cells that we identify as lens fiber cells (Fig 1A-D). The transcription factor *foxe3* is known to be specific to lens epithelial cells (Landgren et al., 2008; Shi et al., 2006), and in our single cell data it was expressed in two cell clusters (Fig. 1A, E), suggesting a difference in global gene expression among lens epithelial cells. These two separate epithelial clusters contained either 1 day post fertilization (dpf) or combined 2 and 5 dpf cells (Fig. 1F and G). There were no apparent differences in clustering between fiber cells at 1, 2 and 5 dpf (Fig. 1F and G).

The shift in epithelial gene expression between 1 and 2 dpf coincides with the end of lens delamination from the surface ectoderm, transition of the epithelium to a simple cuboidal layer extending towards the posterior pole of the lens, and the initiation of secondary fiber cell formation (Greiling and Clark, 2009). To examine this transition, we identified preferentially expressed genes in the two lens epithelial cell clusters (those with adjusted p-values ≤ 0.05) and submitted them to gene ontology (GO) term analysis using GOrilla (Eden et al., 2009; 2007). Epithelial cells at 1 dpf were enriched in genes related to 10 GO terms associated with protein translation, ribosome biogenesis, and embryonic development (Supplemental table 1). Epithelial cells at 2 and 5 dpf had shifted to 19 GO terms dominated by membrane transport and ATP metabolism (Supplemental table 1). We did not find any noticeable difference in gene expression between fiber cells at 1, 2 and 5 dpf.

3.2. Lens crystallin genes show diverse expression patterns

The vertebrate lens includes a diverse group of highly expressed proteins collectively called crystallins that provide the refractive index and transparency required to transmit and focus light onto the retina. Some of these proteins likely also play non-optical roles both within and outside the lens. Our single cell gene expression data provide novel information on the timing and location of crystallin gene expression. These data show how expression has diverged in duplicated crystallin genes, especially in the large number of γ-crystallins.

Zebrafish produce one additional α-crystallin protein compared to mammals because of a duplicated αB-crystallin gene (*cryaba* and *cryabb*) (Smith et al., 2006). There is only one identified zebrafish αA-crystallin gene (*cryaa*) (Runkle et al., 2002). Our scRNA-Seq data showed that zebrafish *cryaa* was expressed almost exclusively in 2 and 5 dpf lens fiber cells (Fig. 2A-B, G). The zebrafish αB-crystallin gene with lens-restricted expression in adults, *cryaba* (Posner et

al., 1999), was expressed in a small number of cells located in clusters previously identified (Farnsworth et al., 2019) as lateral line, hindbrain, muscle progenitors and olfactory tissue (Fig. 2C). Expression in each tissue was temporally restricted, with the three hindbrain cells expressing only at 1 dpf, the one olfactory cell at 2 dpf, and the four lateral line (Fig. 2D) and two muscle progenitor cells at 5 dpf. *cryaba* expression was detected in one 5 dpf lens epithelial cell but in no fiber cells. While we report this very restricted *cryaba* expression here, the small number of cells in each tissue with identified transcript may be a technical artifact. The expression of *cryabb* was detected in more cells, with a large cluster identified as hair cells derived from 2 and 5 dpf and a large cluster of otic capsule cells from 5 dpf (Fig. 2E, F). We conclude that *cryaa* expression was lens-focused, and the expression of the two αB-crystallin genes was almost entirely restricted to non-lens tissues during these stages (Fig. 2H).

Little is known about the specific cell types that express zebrafish β-crystallins. We detected expression for 12 of the 13 known zebrafish β-crystallins in our single cell RNA-Seq data, along with four members of the βγ-superfamily (Table 1). Our single cell data match well with publicly available zebrafish developmental whole-embryo bulk-RNA-Seq data (EBI). For example, the EBI bulk-RNA-seq data indicate initial expression for all 13 β-crystallins between 4.3 and 48 hpf, all within the 5 dpf timespan for our single cell data. The one βγ-crystallin that was not found expressed in whole embryos (*crybg2*) was also not identified in our single-cell atlas. Expression of the β-crystallins fit three general patterns: a) expression restricted to fiber cells (multiple βA- and βB-crystallins: Fig. 3A), b) expression primarily in fiber cells but with some expression in epithelium (multiple βA- and βB-crystallins: Fig. 3B) and c) expression outside the lens (for example, *crybb3* was expressed in cell clusters associated with pharyngeal and olfactory epithelium: Fig. 3C). The EBI bulk-RNA-Seq showed that *crybb2* was the latest expressed β-crystallin with mRNA not detected until 48 hpf (EBI). Our single-cell atlas showed *crybb2* expression in several cranial ganglion cells and several lens fiber cells, all of which were 2 and 5 dpf (Fig. 3D). The non-lens crystallin paralogs *crybg1a* and *crybg1b* had very similar expression in periderm, gills, pharyngeal epithelium, otic capsule and hair cells (Fig. 3E-F). One noticeable difference between the two paralogs was greater expression of *crybg1a* in endothelia and of *crybg1b* in notochord cells. Our single cell atlas found *crybgx* expression exclusively in lens fiber cells starting at 2 dpf, and no *crybg2* expression through 5 dpf (Table 1), similar to the whole embryo RNA-Seq data (EBI).

The γ-crystallins are the largest family of zebrafish lens crystallins due to large gene duplications in the aquatic-specific γM-crystallin group, along with additional duplications in the γN and γS groups. The large number of methionine rich γM-crystallins (34 annotated in zebrafish) has been hypothesized to maintain lens transparency when higher protein densities are required for light refraction in aquatic environments (Kiss and C.-H. C. Cheng, 2008). Zebrafish lack a fourth class found in mammals, the γA-Fs (Wistow et al., 2005). The EBI bulk-RNA-Seq data indicated that unlike a number of β-crystallins that are expressed prior to 20 hpf, the first γM-crystallins are not expressed until 24 hpf (Table 2; EBI). Eleven γM-crystallins were not detected in the EBI bulk-RNA-Seq in whole embryos through 5 dpf (Table 2; EBI). These γM-crystallins whose expression was lacking from the EBI bulk-RNA-Seq data were generally also not detected in our single-cell atlas, indicating concordance with our single-cell atlas. Unlike the expansion of gene expression for several β-crystallin genes into the lens epithelium and outside of the lens, γM-crystallin gene expression was almost entirely restricted to lens fiber cells (Fig. 3G). The only

exception was *crygm3*, which showed expression in some endothelial, intestinal epithelia, and liver cells (Fig. 3H).

Zebrafish express two γN-crystallin genes compared to the single mammalian copy, with *crygn2* expressed earlier in development than *crygn1* (EBI) and its product more common in insoluble protein fractions compared to *crygn1*, possibly due to greater interaction with plasma membranes (Posner et al., 2008; Wistow et al., 2005). Our single-cell atlas confirms stronger expression of *crygn2* in the lens through 5 dpf and shows weak expression outside the lens across the entire embryo (Fig. 3I,J). Furthermore, *crygn1* appears more lens specific with some minor expression in cells of the liver and intestine (Fig. 3I). The function of γN-crystallin is not known. It is expressed in mouse but not in human, and in zebrafish the divergence in expression pattern between the two genes suggests a divergence in functions.

The single copy γS-crystallin protein in mammals is expressed as four proteins in zebrafish. The EBI bulk-RNA-seq data showed no expression of *crygs3* and *crygs4* through 5 dpf, but found low expression of *crygs1* at 48 and 72 hpf and low transient expression of *crygs2* at only 4.3 hpf (EBI). Our single-cell atlas shows no expression of any γS-crystallin in the lens through 5 dpf, consistent with our earlier proteomic analysis that could not identify these proteins through 10 dpf (Wages et al., 2013). However, our single-cell atlas did identify *crygs1* transcripts in skeletal muscle cells (Fig. 3K) and *crygs2* in kidney, brain and olfactory epithelia (Fig. 3L). γS-crystallins are expressed only outside the lens prior to 5 dpf.

3.3. Zebrafish and mammals share a similar set of lens cytoskeletal and membrane proteins

The vertebrate lens contains a unique set of cytoskeletal and integral membrane proteins (FitzGerald, 1990; FitzGerald and Gottlieb, 1989; Gounari et al., 1993; Hess et al., 1996). The zebrafish lens largely includes the same proteins that are well characterized in mammals. The vertebrate lens expresses two unique intermediate filament protein genes, *bfsp1* (*filensin*) and *bfsp2* (*CP49/phakinin*). Both are found as single copy genes in the zebrafish and their transcripts can be detected starting at 19 hpf in whole embryos (Table 3;EBI). In our single-cell atlas, *bfsp1* is expressed in a small number of lens fiber and epithelial cells while *bfsp2* is restricted to a fairly small population of lens fiber cells (Fig. 4A-D; Table 3). We found five spectrin genes in the single-cell atlas, with *sptan1* most highly expressed in the lens and fiber cell preferred (Fig. 4E). In addition, *sptbn1*, *sptbn2*, and *sptbn5* were also detected in lens while also broadly expressed across the embryo in multiple non-lens tissues, while *sptb* was specific to red blood cells (data not shown). There is one copy of the vimentin gene (*vim*) in zebrafish and it is abundant in the nervous system, but also found in the lens with more abundance in fiber cells (Fig. 4F).

The zebrafish lens for the most part expresses the same contingent of cytoskeletal associated plasma membrane proteins as well studied mammalian models. For example, the zebrafish lens expresses the same single cadherin, *cdh2*, which we found broadly expressed across many cell types, and in lens was found evenly distributed in epithelial and fiber cells (Fig. 4G). Four gap junction/connexin genes were found expressed in the lens. One of them, *cx23* (*gje1*) was only expressed in lens cells, and primarily in fiber cells with some weaker expression in 24 hpf epithelial cells (Fig. 4H). Interestingly, cx23 was not expressed in 2 and 5 dpf epithelial cells. We next analyzed *cx43* expression, which was broadly expressed across the embryo, and in the lens

was primarily found in epithelial cells (Fig. 4I). In contrast, *gja3* was expressed almost exclusively in a small number of lens fiber cells (Fig. 4J). The single copy mammalian gene *Gja8* is duplicated in zebrafish with both paralogs almost exclusively expressed in the lens (Fig. 4K, L). The EBI bulk-RNA-seq data showed some temporal differences in expression between *gja8a/*b paralogs: *gja8b* is expressed earlier at 19 hpf compared to 30 hpf for *gja8a* (EBI). Our single-cell atlas also found earlier expression of *gja8b,* in 1, 2 and 5 dpf lens epithelial cells, while *gja8a* was mostly lacking in 1 dpf epithelial cells. Both paralogs were found in lens fiber cells, but *gja8b* was more abundant (Fig. 4K, L).

3.4. Spatial and temporal expression dynamics of genes controlling lens development

A number of transcription factors (TFs) involved in vertebrate lens development have been identified, primarily from studies on mammals and birds (see for review Cvekl and Xin Zhang, 2017). While some of these regulatory genes have also been analyzed in zebrafish (for example (Shi et al., 2006; 2005; Swan et al., 2012)), there has been no comprehensive comparison of TF expression patterns between zebrafish and mammalian species. Interestingly, a number of the known mammalian lens development-related regulatory genes exist as paralogs in zebrafish, conserved after the teleost genome duplication. Our single-cell RNA-Seq atlas provides a foundation for examining the expression patterns of known lens-related regulatory genes, possible divergence in the function of paralogs, and the identification of novel regulators.

The best studied lens development regulatory proteins in zebrafish are those that control fiber cell differentiation, such as *foxe3*, *pitx3* and *hsf4* (Gao et al., 2017; Shi et al., 2006; 2009). Our single cell data showed *foxe3* enriched in lens epithelia (Fig. 5A) while *pitx3* was expressed in both epithelia and fiber cells (Fig. 5B). In the EBI bulk-RNA-Seq data, *hsf4* expression onset is relatively late, beginning at 36 hpf (EBI). Our single-cell atlas identified *hsf4* in a relatively small number of lens fiber cells from 2 or 5 dpf zebrafish (Fig. 5C). The homeobox gene *Prox1* is known to regulate fiber cell differentiation in the mammalian lens (Audette et al., 2016; Wigle et al., 1999). Zebrafish have two paralogs of this gene. Our single cell data show strong *prox1a* expression in lens fiber cells and 24 hpf epithelial cells (Fig. 5D), but almost no lens *prox1b* expression. No study has examined the effects of *prox1a* loss on lens development, although it was shown to play a role in the developmental of lateral line hair cells (Pistocchi et al., 2009). A final TF shown to directly regulate fiber cell differentiation and lens crystallin expression in mammals is *Sox1* (Nishiguchi et al., 1998). A mouse *Sox1* knockout misexpressed *Pax6* during lens development, suggesting a genetic interaction between these transcription factors in the lens (Donner et al., 2007). Zebrafish contain *sox1a/b* paralogs and while their role in neurogenesis has been investigated (Gerber et al., 2019), nothing is known about their role in lens development. Our single-cell atlas showed *sox1a/b* paralogs to have divergent temporal expression in lens epithelial cells, with *sox1a* restricted mostly to 1 dpf epithelial cells while *sox1b* was expressed in 2 and 5 dpf epithelial cells (Fig. 5E, F). Both paralogs were expressed in lens fiber cells.

The large Maf family of transcription factors are well known regulators of lens development (Cui et al., 2004; Kim et al., 1999; Reza and Yasuda, 2004; Takeuchi et al., 2009; Yang et al., 2004). Vertebrates express four classes of large Mafs with mammals expressing one gene for each class (Coolen et al., 2005). The most critical large Maf in mammalian lens development, c-Maf, occurs

as two paralogs in zebrafish, *mafa* and *mafb*, which were both expressed in lens in our single cell data (Fig. 5G, H). Another large Maf, MafB, also occurs as paralogs in zebrafish. But in this case only one of the paralogs, *mafba*, is expressed in lens (Fig. 5I). The last two large Mafs, MafA and NRL occur as single orthologs in zebrafish, with *nrl* expressed in lens but *mafaa* only outside the lens (data not shown).

Almost all of the transcription factors categorized by Cvekl and Zhang (Cvekl and Xin Zhang, 2017) as being involved in mammalian *pax6* signaling, lens vesicle separation, lens cell survival, cell proliferation and differentiation were expressed in lens cells in our single-cell atlas (Fig. 6A,B). The only exceptions, *pknox1.1* and *hif1aa,* have paralog pairs that were expressed in lens. These data indicate a broad retention of lens expression of mammalian-related lens development regulators. Our data also showed that most of these transcription factor genes, unlike the lens-focused expression of genes that directly regulate crystallin expression, were expressed in a broad range of extralenticular tissues (Supp fig. 1). The expression of lens development-related transcription factor genes within the lens varied in both intensity and their relative expression within epithelial or fiber cells (Fig. 6). Most were more strongly expressed in epithelial cells, although *sox2* and *prox1a* were notable exceptions (compare Fig. 6A to 6B).

While studies of lens development have focused primarily on the role of transcriptional control of gene expression, recent studies indicate that proteins bind to the untranslated regions of lens mRNAs to modify their expression. Of the over 1600 RNA-binding proteins (RBPs) encoded in the mammalian genome only four have been implicated so far in lens function: Tdrd7 (Lachke et al., 2011), Caprin2 (Dash et al., 2015), Celf1 (Siddam et al., 2018) and Rbm24 (Dash et al., 2020; Grifone et al., 2018). CRISPR-mediated mutation of zebrafish *rbm24a* causes reduced or absent eyes (Grifone et al., 2018). This phenotype may result from the role *rbm24a* plays in regulating lens crystallin mRNA stability (Shao et al., 2020). A recent study showed that morpholino-induced translation blocking of *celf1* causes micropthalmia and visibly cloudy lenses by 4 dpf (Siddam et al., 2018). Our scRNA-Seq data confirm that both genes are strongly expressed in lens (Fig. 5J, K). The functions of *tdrd7a*, *tdrd7b* and *caprin2* have not been examined in zebrafish. Our data show that *tdrd7b*, but not *tdrd7a*, is expressed in lens (Fig. 5L) and *caprin2* is only expressed in a few lens cells (data not shown).

3.5. Identification of possible novel regulators of lens development

Transcriptomic analysis of the mouse lens using RNA-Seq and microarray data has helped identify a number of novel genes involved in lens development (Kakrana et al., 2017; Siddam et al., 2018; Srivastava et al., 2017). These studies hypothesized that genes with lens-specific expression are important for lens function and assessed that function using mutant analysis. This approach has successfully identified a number of genes that regulate lens development, including *Celf1*, *Tdrd7*, *Mafg* and *Mafk*. We hypothesized that genes that are important for lens development and function would be strongly enriched in the lens within our scRNA-seq data. To find enriched genes in the developing zebrafish lens, we ranked genes by their preferred expression in either lens epithelia or fiber cells (Supplemental tables 2 and 3). These rank lists include multiple genes already known to be important in lens development, but also identify novel genes.

The top 200 lens fiber cell-preferred genes, ranked by expression relative to all cells outside the cluster, included 12 TF genes and six RBP genes. Of the 12 TF genes in the list, seven are already known to play roles in lens development (*etv4*, *mafba*, *mycn*, *pitx3*, *prox1a* and *sox1a*/*sox1b*). An eighth, *prox2*, has been knocked out in mice but no lens phenotype was noted (Nishijima and Ohtoshi, 2006). Strong lens expression of *prox2* in zebrafish embryos was also identified by whole mount in situ hybridization (Pistocchi et al., 2008). Morpholino-induced depletion of a ninth gene on our list, *hmx4*, in zebrafish led to a reduced forebrain and narrowed eye field due to disruption in retinoic acid and Sonic hedgehog signaling, but no lens phenotype was examined (Gongal et al., 2011). However, whole mount in situ hybridization analysis did confirm *hmx4* expression in lens (French et al., 2007; Song and Perkins, 2018). The other three TF genes with strong fiber cell preferred expression (*btf3l4*, *hmx1* and *sox13*) have not been investigated for their role in lens development and provide novel targets for future studies (Fig. 7; Table 4). A SNARE protein gene on the list, *stx3a*, was recently shown to produce a protein isoform in humans that, in epithelial tissue culture, could regulate gene expression through interactions with the transcription factor ETV4 (Giovannone et al., 2018).

Our list of fiber cell enriched genes also included seven RNA-binding proteins (RBPs), including two that were previously implicated in lens development, *celf1* and *rbm24a* (Fig. 5G and H; Supp Table 2). The other five have no previously described role in lens development or function (Fig. 7). One of them, *pabpc4*, regulates mRNA stability by binding to poly(A) regions, similar to other known lens-related RBPs (D. Liu et al., 2012). Two others (*endou2* and *endouc*) are members of the EndoU ribonuclease family whose members play roles in regulating ER structure, RNA degradation and cell survival (Lee et al., 2017). Available whole mount in-situ data show *endou2* to be expressed strongly in lens by 24 hpf (Thisse and Thisse, 2004), matching our scRNA-Seq data. No role for the EndoU gene family has been previously described for the lens. Another RBP on our list of fiber cell preferred genes, *ssb*, is known to bind to nascent RNA transcripts as well as the poly(A) tail of mRNAs (Vinayak et al., 2018). A final fiber cell preferred RBP, C1qpb, binds mRNA primarily in mitochondria and is linked to several types of cancer and regulated cancer cell chemotaxis and metastasis (Xiaofang Zhang et al., 2013).

A list of the top 200 lens epithelial cluster enriched genes included a large number related to transcriptional machinery, ribosome function and protein synthesis (Supp table 3). There were also multiple genes involved in ATP synthesis and amino acid transport. One amino acid transporter gene, *slc7a11*, showed highly specific expression in lens epithelium and a transketelose (*tkta*) was highly expressed in 2 and 5 dpf epithelial cells (Fig. 7; Table 4). A gene involved in arginine biosynthesis, *otc*, was also highly lens specific and found in both epithelial and fiber cells (Fig. 7; Table 4). We identified five crystallin genes that were enriched in the putative lens epithelial cell cluster, including β-crystallins as well as *crygmx*. Sixteen transcription factor (TF) genes were enriched in lens epithelial cell clusters. Nine of them (*btf3l4*, *hmx1*, *hmx4*, *mafba*, *pitx3*, *prox2*, *sox1a*, *sox1b, stx3a*) were also enriched in fiber cells (Supp tables 2 and 3). Nine were previously implicated in lens development (*foxe3*, *mafba*, *mycb*, *pax6a*, *pax6b*, *pitx3*, *six3b, sox1a, sox1b*), further showing how our single cell data can highlight functionally important genes. Two additional TF genes (*id3*, *nr2f6b*) showed lens epithelial-preferred expression but have no previously described lens function (Fig. 7; Table 4). There were six RNA binding proteins (RBPs) in the top 200 epithelial gene list, including five that we also found to be enriched in fiber cells (*c1qbp*, *celf1*, *endouc*, *rbm24a*, *ssb*). The sixth RBP gene

serbp1a, was expressed in both epithelial and fiber cells, as well as many tissues outside of the lens and has no known lens function. Lastly, our list of lens epithelial-preferred genes included *bmp5*. While BMP signaling is known to be involved in lens development (Belecky-Adams et al., 2002; Cvekl and Duncan, 2007) *bmp5* specific function has not previously been identified in the lens.

4. Discussion

The data presented here are the first single cell level analysis of lens-related gene expression in any species. Lens cells clustered by UMAP non-linear dimensional reduction were identifiable as either fiber or epithelial cells based on the expression of previously identified marker genes. Cell clustering indicated a global shift in gene expression of epithelial cells between 1 and 2 dpf coinciding with secondary fiber cell differentiation. There was no noticeable change in fiber cell global gene expression with age. Our data also provide a detailed description of lens crystallin gene spatiotemporal expression, most significantly clarifying the early expression of the three small heat shock α-crystallin proteins and identifying various patterns of expression in the diverse β- and γ-crystallins. These data provide a new perspective on long-standing questions about lens crystallin evolution and function and possible roles for these proteins outside of the vertebrate lens. Examination of lens cell membrane-associated and cytoskeletal protein genes indicate strong levels of conservation with previously described orthologs from mammalian species, although with some divergence in duplicated genes. Analysis of known lens-related transcription factor genes shows general conservation between mammals and zebrafish, but again with divergence between duplicated paralogs. These single cell expression data identify multiple additional genes with lens preferred expression but with no currently described lens function, including several transcription factor and RNA-binding proteins. In total, our analysis provides a temporospatial atlas of known lens-related genes during early development while also highlighting a number of potential targets for future study.

An important goal of single cell expression analyses is to identify genes with previously undescribed functions in specific tissues. Our lists of lens epithelial and fiber cell preferred genes include at least 16 with no known lens function (Fig. 7; Table 4). Identification of lens preferred expression in the mouse has been a productive approach for identifying new genes with lens related function (Kakrana et al., 2017; Siddam et al., 2018; Srivastava et al., 2017). The majority of the novel genes we identified were for transcription factors and RNA-binding proteins. Follow-up studies on these genes using CRISPR/Cas9 gene targeting could help deepen our understanding of the gene regulatory mechanisms that influence lens differentiation. Identification of novel RNA-binding proteins helping regulate lens development would be particularly impactful as only four such genes have so far been identified for the lens. Several additional genes that could be targeted for study coded for metabolic proteins, possibly leading to new information about lens physiology. The inclusion in our epithelial and fiber cell preferred gene expression lists of multiple genes just recently shown to play a role in lens development, such as the RNA-binding proteins *celf1* and *rbm24a*, supports the hypothesis that the novel genes in our lists are good targets for future gene targeting experiments.

Comparisons between the expression of mammalian and zebrafish lens developmental transcription factors (TFs) are instructive in better understanding their function. Here we

summarize three observations about our TF expression data. First, almost every TF known to be involved in mammalian lens development was expressed at some level in zebrafish lens cells. This overall retention of lens expression in the zebrafish orthologs of mammalian lens regulatory genes fits the hypothesis that the mechanics of lens development are conserved within vertebrates. Second, most mammalian lens development-related transcription factors do not show unique or elevated expression in zebrafish lens cells, but instead typically exhibit expression in diverse tissues. For example, almost all of the genes involved in the *pax6* lens regulatory network, including the two zebrafish *pax6* paralogs, show this pattern (Supp fig. 1). Similarly widespread expression was seen in most genes involved in lens cell survival, lens vesicle separation, and cell proliferation/differentiation. While few studies have directly examined whether these genes are involved in zebrafish lens development, it is notable that so many are expressed in lens, while at the same time exhibiting such diverse expression in non-lenticular tissues. Many of the genes known to directly regulate lens crystallin expression (eg. *hsf4*, *prox1a*, *sox1a* and *sox1b)* were more specifically expressed in lens cells (Figure 6; Supp Fig 1). The strong levels of lens cell specific expression in this crystallin regulatory group suggests two possible hypotheses: first, crystallin regulation is highly conserved in the vertebrate lens and this evolutionary conservation is reflected in strong gene expression. However, a second possible hypothesis is simply that the large concentrations of crystallin proteins generated by lens cells requires large concentrations of their regulatory transcription factors.

A third and final observation about lens TFs is that the expression patterns of paralogs are often only partially overlapping. For example, while the two *pax6* paralogs have largely overlapping expression in lens, retina, brain and olfactory epithelia, *pax6b* showed unique expression in pancreatic cells (Supp Fig 1). Furthermore, *six3a* and *six3b* showed overlapping expression in forebrain and olfactory cells, but *six3b* was more highly expressed in lens and retina (Supp fig. 1). Most interesting for the study of lens development is the divergent expression between lens epithelial and fiber cells of several paralog pairs. The greater expression of *sox1b* in fiber cells and 2-5 day epithelial cells compared to *sox1a* may reflect subfunctionalization that could be explored through future knockout studies (Fig. 6). A recent study showed co-expression of both zebrafish *sox1* paralogs in brain, lens, otic vesicle and spinal cord by whole embryo in situ hybridization (Gerber et al., 2019). This result is in agreement with the *sox1a/b* expression pattern observed from our single-cell atlas*.* Relative shifts in expression were also found in the paralog pair *myca*/*mycb*, the orthologs to mammalian C-Myc. For example, while *myca* expression is slightly higher in fiber cells its paralog *mycb* shows a shift to epithelial expression (Fig 6). While they are expressed at lower levels overall, there is similar lens cell expression divergence in the paralogs for *tfap2a/b*, *otx2a/b* and *meis1a/b*. We do not know if there are any functional consequences to these shifts in expression. In addition, the temporal dynamics of transcription factor expression within lens epithelial and fiber cells found in our single-cell atlas are striking. For example, *pax6a* and *pax6b* show increasing expression over time (1-5 dpf) in lens epithelial cells, but not in fiber cells (compare Fig. 6A to 6B).

The developmental time points we examined surround the key period when secondary lens fiber cells differentiate between 1 and 2 dpf and the epithelium becomes a simple cuboidal layer. During this period we found a shift in lens epithelial cell gene expression from intense protein production to membrane transport and ATP generation (Supp Table 1). This shift may reflect the epithelium's more terminal role in metabolically sustaining the growing mass of denucleated

fiber cells. We do not believe that this early developmental physiological shift has been identified by previous studies. It will be interesting to see if this change in epithelial cell biology is conserved in other vertebrate species as additional single cell analyses are published. Lens fiber cells undergo large morphological changes during the first five days of zebrafish development as secondary fiber cells are added between 1 and 2 dpf and fiber cell denucleation is completed around 3 dpf. However, we did not observe distinct clustering between fiber cells at the three developmental ages examined in our study. Since new fiber cells continue to be produced during lens growth at the stages analyzed in this study, each larva, and thus each sample, may contain both newly born and more differentiated fiber cells that could be transcriptionally similar. It is also possible that strong attachments between maturing fiber cells may bias against their ability to be efficiently profiled using our dissociation methods and scRNA-seq library preparation. In total these data suggest that epithelial cells undergo a significant shift in activity during the first 5 days of zebrafish lens development that is not paralleled in the fiber cells.

A substantial literature on the structure, function and expression of lens crystallin proteins has provided a rich understanding of their roles in lens development, function and disease (Andley, 2007). We believe that our single cell expression dataset provides a valuable, novel perspective on lens crystallins. The most studied lens proteins have been the α-crystallins due to their hypothesized role in maintaining lens transparency and evolutionary relationship with small heat shock proteins (Horwitz, 1992). Our single cell data mirror previous RT-qPCR and RNA-Seq data that indicated an increase in *cryaa* expression starting at 2 dpf (Posner et al., 2017). *Cryaa* expression was detected at an earlier stage in embryonic mouse lens using *in situ* hybridization, where transcripts were found in the lens cup by E10 (Robinson and Overbeek, 1996). This stage is equivalent to approximately 20 hpf in the zebrafish. While we have not captured this stage in our single cell data, publically available whole embryo RNA-Seq data do not show *cryaa* expression at this stage either, suggesting that the zebrafish may delay *cryaa* expression relative to mouse. Later in development mouse *Cryaa* expression becomes most prevalent in fiber cells (Robinson and Overbeek, 1996), similar to what we found in our zebrafish single cell data.

There are conflicting data in the literature about the early expression pattern of the two zebrafish αB-crystallins (see (Posner et al., 2017) for discussion). Briefly, RT-PCR and qPCR of whole embryos did not detect expression of *cryabb* through 78 hpf and 5 dpf, respectively, while qPCR did detect *cryaba* starting at 12 hpf (Elicker and Hutson, 2007; Mao and Shelden, 2006). We previously showed a similar pattern by qPCR, with *cryaba* expression increasing between 4 and 5 dpf but minimal expression of *cryabb* (Posner et al., 2017). Another recent study detected both α B-crystallin genes starting at 24 hpf (Zou et al., 2015). Our single cell RNA-Seq data help address these discrepancies by identifying the embryonic cell types expressing each gene. No *cryabb* mRNA was detected in lens cells through 5 dpf and only one lens cell, a 5 dpf epithelial cell, expressed *cryaba*. Our data suggest that both αB-crystallin mRNAs are essentially non-existent in the lens through 5 days of development. However, both genes were expressed at varying levels outside the lens. What functional role, if any, these proteins play in those cells is not known. In adult zebrafish *cryaba* becomes mostly lens-specific while *cryabb* is expressed broadly, similar to the mammalian ortholog (Smith et al., 2006). The timing of these ontogenetic shifts in expression, and when exactly both αB-crystallin genes are expressed in lens, will require examination of older developmental stages. It is interesting that in embryos, as in adults, *cryabb*

has the most diverse tissue expression of the three α-crystallin genes. Neither zebrafish αB-crystallin gene shares the same expression pattern as mouse αB-crystallin, which was detected in the lens placode of mouse embryos and became preferentially expressed in the lens epithelium as primary lens fiber cells differentiated (Robinson and Overbeek, 1996), a pattern quite different from that of the zebrafish.

It is possible that αB-crystallin proteins may be accumulating in lens cells without detectable levels of mRNA, or that we are simply failing to detect their mRNAs (a false negative), but considering that we were able to detect mRNA for a number of other lens crystallins at these same ages this interpretation seems unlikely. A previous mass spectrometry examination of zebrafish lenses found neither αB-crystallin in the lens through three weeks of age, but abundant protein at 6 months of age (Greiling et al., 2009). While future proteomic examination of embryonic zebrafish lenses would provide additional data regarding the timing of αB-crystallin accumulation in the lens, it appears that these proteins are not present during early development. We previously showed that the proximate promoter regions for the two zebrafish αB-crystallins did not efficiently drive GFP expression in the embryonic lens, mirroring our new scRNA-seq data (Posner et al., 2017). However, that study also showed that both zebrafish αB-crystallin promoters drove GFP expression in skeletal muscle and notochord in embryos at 2 dpf, a result that is not reflected in our single cell data. It is possible that more distal enhancer regions alter the expression of these genes, providing a caution when interpreting analysis of isolated promoter regions.

Much less is known about the function of β- and γ-crystallins in the vertebrate lens. Most work on this group has focused on the impact of mutations, post-translational modifications and protein aging in cataract development (Lampi et al., 2014), although some studies have examined the roles of β-crystallins in retinal astrocytes and in phagocytosis within the retinal pigmented epithelium (Sinha et al., 2008; 2012; Zigler et al., 2011). Zebrafish have a similar contingent of β-crystallins compared to mammals, but with added genes due to duplications (Wistow et al., 2005). The contingent of zebrafish γ-crystallins is quite different from mammals due to the presence of the aquatic lens γM-crystallins and loss of the ϒA-F-crystallins. No study has presented comprehensive data on the spatial expression of β- and γ-crystallins in any species.

Mammalian lenses express eight β-crystallin proteins using seven genes (Wistow, 2012). Zebrafish have expanded this family to 13 genes, seven βA- and six βB-crystallin genes. The largest expansion is of mammalian βA1-crystallin. While this one gene produces two different proteins in mammals, zebrafish have four βA1 genes encoding four known proteins. The two paralogs expressed earliest in development (*cryba1l1* and *cryba1b*) also showed more extensive lens expression than the other two β A1-crystallins, extending from fiber cells into epithelia. Future work will need to explore whether these differences in expression are connected to variation in function. The largely overlapping expression of the two zebrafish βA2 paralogs does not suggest functional divergence, but this question would also need to be examined experimentally. The single βB1 mammalian crystallin is found as four genes in zebrafish. Onset of expression as shown in whole embryo RNA-Seq data is quite divergent between the four genes, which may suggest unique roles of these paralogs. The spatial range of expression from fiber

cells into epithelial cells also differed. The single βB2 gene showed limited expression in lens fiber cells as well as expression in a few cells identified as cranial ganglia. The only βA- or βB-crystallin with significant expression outside the lens was βB3-crystallin. This observation is interesting considering that mammalian *Crybb3* is also unique among β -crystallins in being expressed transiently within the nuclei of differentiating lens fiber cells (Limi et al., 2019).

No studies have examined the role of the fish specific β-crystallin *crybgx*, and its very limited expression in our dataset (only found in two lens fiber cells) does not suggest a vital role in lens development. Zebrafish contain three genes for the non-lenticular members of the β-crystallin family. The encoded proteins are often not referred to as crystallins since they are not produced in the lens, and in humans are known as the AIM1 and AIM1-like proteins. The two zebrafish paralogs for human *AIM1*, *crybg1a* and *crybg1b*, showed largely overlapping expression outside the lens, again suggesting conservation of function between paralogs.

Despite the large number of zebrafish γM-crystallin genes, most of them showed remarkably similar expression, with onset around 30 hpf when secondary fiber cells begin to differentiate. These data showing late and lens specific use support the hypothesis that the role of γM-crystallins is restricted to a building material allowing for close packing of proteins in a high density aquatic lens. According to this hypothesis, diversity in protein shape provides for the polydispersity that helps prevent protein aggregation and opacity at high density. While α-crystallins achieve this polydispersity by forming a range of aggregate sizes (Aquilina et al., 2005) γ-crystallins form uniform dimers, so multiple proteins may be needed to achieve this same effect in the aquatic vertebrate lens. While the diversity of γM-crystallins is not know from the sarcopterygian fishes that are more closely related to tetrapods, it is possible that this entire class of ϒ-crystallins was lost during the transition of vertebrates to land.

Possible roles played by the other two classes of ϒ-crystallins remain unclear and a potentially interesting open question. The lack of mRNA for any of the four zebrafish γS-crystallins in the lens through 5 dpf matches previous proteomics studies (Greiling et al., 2009; Wages et al., 2013). Together these results suggest that ϒS-crystallins play no role in lens development, but do become abundant members of the adult lens proteome (Posner et al., 2008). While their role in the lens is yet to be determined, the early expression of *crygs1* in skeletal muscle and *crygs2* in olfactory and kidney tissue may indicate non-ocular developmental roles. Further study of these two genes may help us understand the roles of the single copy γS-crystallin in mammals. Likewise, the two zebrafish paralogs of the single copy mammalian γN-crystallin have diverged in timing of expression onset and extent of lens expression. The function of these proteins in any species is largely unknown, making zebrafish an excellent model for exploring their biological roles. It is worth noting that many of the βA-, βB- and γM-crystallins showed very weak expression in a wide range of embryonic tissues. This pattern was not seen in α-crystallins, suggesting that it is specific to these crystallin classes. Whether this weak background expression reflects any functional importance in these tissues is not known.

Unlike the large numbers of paralog pairs in the crystallins and lens regulatory genes, zebrafish membrane associated and cytoskeletal proteins occur mostly as one-to-one orthologs with mammals. The only exception is the presence of two paralogs for *gja8*, which did differ in

expression and could possibly be useful for studying this protein. The single copy *gja3* (AKA cx48.5) provides an interesting test of how well different experimental approaches agree when analyzing gene expression and function. A previous study identified *gja3* mRNA by RT-PCR in 4 dpf zebrafish lens and heart, while whole-mount in-situ hybridization found mRNA in lens and transiently in otic vesicles, but not in heart (S. Cheng et al., 2004). Our scRNA-Seq data found *gja3* expressed in lens fiber cells, with expression in some cells of the periderm and gills, but not in heart or otic vesicles (Fig. 4J, some data not shown). Interestingly our single cell data and the previous study's in-situ hybridizations both showed that lens *gja3* expression is restricted to fiber cells. Morpholino-induced reduction of *gja3* expression produced lens abnormalities and cardiovascular defects in 3 dpf embryos (S. Cheng et al., 2004). Further comparisons of data resulting from different gene expression analysis approaches (ie. RT-PCR, RT-qPCR, RNA-Seq, scRNA-Seq, promoter analysis) combined with functional testing of gene function using translation blocking and knockout tools will help to clarify how well gene expression levels predict functional importance in a specific tissue. Throughout our results section we have attempted, when possible, to compare our single cell RNA-Seq data to published data from in situ hybridization qPCR and bulk RNA-Seq experiments to determine the level of concordance between techniques. We generally found that these data match, but recommend that any follow up analyses of gene expression patterns described in our single cell dataset should be confirmed by in situ hybridization analysis. In the case of *gja3* it is interesting that some techniques showed similar expression in some tissues, but not others. The demonstration that reduction in *gja3* expression affects heart development is not predicted by expression patterns resulting from in situ hybridization and single cell RNA-Seq. It is worth noting that single cell RNA-Seq can lead to false negatives if transcripts are not captures, but false positives are unlikely. Therefore, identification of novel genes not previously shown to have lens preferred expression is likely a useful approach for selecting targets for future analysis of lens development and function.

We noticed a number of lens-related genes that shared expression with tissues derived from the preplacodal region. This vertebrate embryonic tissue is a shared developmental origin for the lens and hair cell-containing sensory organs like the inner ear, olfactory epithelium and lateral line (Bailey and Streit, 2005; Grocott et al., 2012). Examination of the molecular signaling that leads to differentiation of these different sensory structures showed that the default gene expression pattern leads to lens development, with the other structures requiring shifts in the production of signaling molecules (Bailey et al., 2006). Study of vertebrate tissue regeneration also shows shared molecular signaling pathways between the lens and inner ear hair cells (Tsonis, 2007; Tsonis et al., 2007). Both αB-crystallins were expressed in extralenticular tissues of preplacodal origin with *cryaba* expressed primarily in one of these tissues, the lateral line, while *cryabb* was found in lateral line, otic capsule, and other hair cells (Fig. 2H). One β-crystallin, *crybb3*, was expressed in olfactory epithelia, two β/γ-crystallins were expressed in otic capsule and lateral line and one γ-crystallin, *crygs2*, was expressed in olfactory epithelium. Interestingly, all of these crystallins with extralenticular preplacodal tissue expression are either non-existent in lens or expressed at lower levels in the lens. This observation suggests that these genes may be pulled away from the lens function of their other crystallin family members by changes in gene expression signaling into other roles in preplacode derived tissues.

Transcription factor genes have also been shown to share expression in placode-derived tissues. Zebrafish *pitx3* was found to be expressed in placode regions that give rise to the lens, olfactory

epithelium, anterior pituitary and cranial ganglia, but not the otic and lateral line placodes (Zilinski et al., 2005). Our single cell data mostly agree, showing *pitx3* most highly expressed in lens epithelium, lens fiber cells and hypophysis, but little expression in hair cell tissues (Supp fig. 1). However, we saw minimal expression in olfactory tissue. We did find a number of lens-related TF genes expressed in a cell cluster identified as olfactory epithelium/placode at greater levels than their general widespread expression (e.g. *pax6a*, *cebpb*, *meis1a*, *six3a* and *six3b*; Supp fig. 1). The same set of genes showed shared increased expression in lateral line and hair cells (*sox1a*, *sox1b*, *tfap2a*, *cebpb*, *meis1a*, *meis1b*, *sox2* and *mycb*), while several genes showed shared expression in hypophysis (*pitx3*, *tfap2a*, *tfapsb* and *cebpb*). There are opportunities to further explore the connection between signaling molecules leading to diversification of placode-derived tissues.

The work presented here provides an embryo-wide view of the expression of genes linked to lens development and function. We do not know of previous studies in any species that characterize the expression of such diverse families of lens genes at single cell resolution. The data address questions about α-crystallin expression and provide the first systemic analysis of zebrafish β- and γ-crystallin expression. The large number of β- and γ-crystallin paralog genes, and the presence of the unique γM-crystallins in the fish lens, allow for a new view of the evolution of these gene families in the vertebrate lens. Analysis of cytoskeletal, membrane and transcription factor genes supports the hypothesis of strong conservation between mammal and zebrafish lenses, while pointing to areas where gene duplication in zebrafish may provide opportunities to further dissect the roles of these genes. Lastly, the identification of genes with preferred expression in lens, sometimes specific to one cell type, shows promise for identifying novel regulators of lens expression, including new RNA-binding proteins, whose functions in lens development have only recently been recognized.

Figure and Table Legends

Figure 1. Expression of known lens marker genes identified lens epithelial and fiber cell clusters. Global single cell gene expression in 1, 2 and 5 dpf embryos was analyzed with UMAP to cluster cells with similar expression patterns (A). Two areas in the total atlas (A) that contained lens cells are indicated with dotted lines and arrows. Four genes known to be expressed in either lens fiber cells (*cryba1a* (B); *grifin* (C); *mipa* (D)) or epithelial cells (*foxe3* (E)) were used to identify these cell types (F). The intensity of purple coloration in panels B-E reflects gene expression. Panel G shows the developmental age of all lens cells. Note that insets in panels B-G include the 2 and 5 dpf lens epithelial cells. Gene ontology term analysis of the most expressed genes in 1 dpf compared to 2 and 5 dpf lens epithelial cells showed a shift from enrichment in protein translation, ribosome biogenesis and embryonic development at 1 dpf to membrane transport and ATP metabolism at 2 and 5 dpf (Supp Table 1).

Figure 2. Lens fiber cells upregulated expression of αA-crystallin starting at 2 dpf, but αB-crystallin expression in the lens was almost non-existent through 5 dpf. mRNA for *cryaa* was found almost exclusively in the lens (A), and primarily in fiber cells (B). Expression of *cryaba* was detected primarily in lateral line cells (arrowhead in C and detailed view in D). Expression of *cryabb* was detected at relatively low levels in diverse tissues including otic capsule and hair cells (arrowhead in E and details in F). Temporal expression dynamics of *cryaa* expression in lens fiber cells through 5 dpf (G). Expression of α-crystallin genes in tissues with significant representation (H). For dot plots, the size of each dot reflects the percentage of cells in that tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Figure 3. Spatial expression of the β-crystallin and diverse gamma crystallin gene families. Many, like *cryba1l2* were confined to lens fiber cells (A). Others like *cryba1l1* were most strongly expressed in lens fiber cells but also extended into lens epithelium (B). *crybb3* was unique among β-crystallin genes in lacking lens expression and instead being found in olfactory epithelium, pharyngeal and basal fin cells (C). *crybb2* was found in both lens fiber and cranial ganglion cells (D). The paralogs *crybg1a* and *crybg1b* had mostly overlapping expression, although the former was more abundant in endothelia (E) while the latter was abundant in notochord (F). Almost all γM-crystallin genes with detected expression were restricted to lens fiber cells, as shown for *crygm2d10* (G, arrowhead). The only γM-crystallin gene with extralenticular expression was *crygm3*, which was expressed in endothelial, intestinal epithelia and liver cells (H). Both γN-crystallin genes were expressed primarily in lens fiber cells with some lesser expression in lens epithelial cells (I and J). Both were expressed in many other tissues at weaker levels across the embryos, although this expression was stronger for *crygn2*. Only two γS-crystallin genes were expressed in our dataset. *crygs1* was expressed in a number of skeletal muscle cells (K) while *crygs2* was expressed in kidney, brain and olfactory epithelia (L). Neither γS-crystallin gene was expressed in lens.

Figure 4. The zebrafish lens expresses many of the same cytoskeletal and membrane proteins as mammals, with some interesting divergence between paralogous pairs. Panels A and B indicate cell type and age as a reference (see Fig. 1). Shown are the expression of two intermediate filament genes unique to the vertebrate lens, *bfsp1* (*filensin*) and *bfsp2*

(*CP49/phakinin*) (C and D). Also shown is a spectrin gene (E) and the single vimentin gene (F). Multiple cytoskeletal-associated plasma membrane protein genes are also shown, including one cadherin (G) and two connexin genes (H, I). Three gap junction protein genes are shown, including *gja3* (J) and a paralog pair for *gja8a* and *gja8b* (K, L).

Figure 5. Spatial expression of transcription factors and RNA-binding proteins known to regulate lens development. Expression for the three transcription factors *foxe3* (A), *pitx3* (B) and *hsf4* (C) is shown in the identified lens cell clusters. Inset contains two and five dpf epithelial cells as identified in Figure 1G. These three transcription factors have been shown to take part in shared lens development signaling pathways in both mammalian and zebrafish lens. The mammalian gene *Prox1* is duplicated in zebrafish, but only the paralog *prox1a* is expressed in lens (D). Mammalian *Sox1* also occurs as paralogs in zebrafish, and both are expressed in lens, with *sox1a* showing greater expression in 1 dpf epithelial cells (E) while *sox1b* expression appears to extend into 2 and 5 dpf epithelial cells (F: cells within inset). The expression of three large Maf family genes are shown, including the mammalian c-Maf paralogs *mafa* (G) and *mafb* (H) as well as one mammalian MafB paralog *mafba* (I).Three previously characterized RNA-binding protein genes known to regulate lens development were expressed in lens cells: *rbm24a* (J), *celf1* (K) and *tdrd7b* (L).

Figure 6. Temporal dynamics of transcription factor expression in zebrafish lens epithelial and fiber cells at 1, 2 and 5 dpf. Expression of lens development regulators and zebrafish paralogs for single copy mammalian genes in lens epithelial (A) and fiber cell (B) clusters. Functional categories for transcription factors are noted below each panel. The size of each dot reflects the percentage of cells in that tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Figure 7. Expression of genes with preferred expression in lens fiber cells and epithelia, but with no currently identified lens function. The size of each dot reflects the percentage of cells in each tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Supplemental Figure 1. **Expression levels of known regulators of mammalian lens development across all zebrafish cell types**. This dot plot arranges lens development regulators identified in mammalian studies by function as defined in Cvekl and Zhang (2017). Known zebrafish paralogs for single copy mammalian genes are included. Size of each dot reflects expression levels. Cluster numbers can be connected to hypothesized tissue types using supplemental table 4.

Supplemental Table 1. Gene Ontology (GO) term comparison of 1 dpf and 2-5 dpf lens epithelial cells. Lists of preferentially expressed genes in each cluster were submitted to GOrilla. Shown are the GO term ID and descriptor, gene count for each GO term, p-value returned by GOrilla and the top genes for each term (up to a total of five). Only GO terms with a p-value ≤ 0.05 are shown.

Supplemental Table 2. Top 200 genes expressed in a cell cluster identified as lens fiber cells ranked by their log fold change in expression compared to cells outside the cluster. Shown along with gene name are adjusted p-values, fraction of cells within and outside the lens cluster with detected expression, the ratio of that expression, the general category of each gene's function and whether the gene was also in the top-200 expressed lens epithelial genes.

Supplemental Table 3. Top 200 genes expressed in a cell cluster identified as lens fiber cells ranked by their log fold change in expression compared to cells outside the cluster. Shown along with gene name are adjusted p-values, fraction of cells within and outside the lens cluster with detected expression, the ratio of that expression, the general category of each gene's function and whether the gene was also in the top-200 expressed lens epithelial genes.

Supplemental Table 4. List of hypothesized cluster identifications by number.

Acknowledgements

We would like to thank Adil Hussen for help with the analysis of crystallin gene expression and the Busch-Nentwich lab for providing RNA-seq data. Funding: This work was supported by the NIH National Eye Institute [R15 EY13535] to M.P. and by the NIH Office of the Director [R24 OD026591], NIH National Institute of Neurological Disorders and Stroke [R01 NS105758], and the University of Oregon to A.C.M. Funding sources played no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Citations

Agrawal, S.A., Anand, D., Siddam, A.D., Kakrana, A., Dash, S., Scheiblin, D.A., Dang, C.A., Terrell, A.M., Waters, S.M., Singh, A., Motohashi, H., Yamamoto, M., Lachke, S.A., 2015. Compound mouse mutants of bZIP transcription factors Mafg and Mafk reveal a regulatory network of non-crystallin genes associated with cataract. Hum. Genet. 134, 717–735. doi:10.1007/s00439-015-1554-5

Andley, U.P., 2007. Crystallins in the eye: Function and pathology. Progress in retinal and eye research 26, 78–98. doi:10.1016/j.preteyeres.2006.10.003

Aquilina, J.A., Benesch, J.L.P., Ding, L.L., Yaron, O., Horwitz, J., Robinson, C.V., 2005. Subunit exchange of polydisperse proteins: mass spectrometry reveals consequences of alphaA-crystallin truncation. J Biol Chem 280, 14485–14491. doi:10.1074/jbc.M500135200

Audette, D.S., Anand, D., So, T., Rubenstein, T.B., Lachke, S.A., Lovicu, F.J., Duncan, M.K., 2016. Prox1 and fibroblast growth factor receptors form a novel regulatory loop controlling lens fiber differentiation and gene expression. Development 143, 318–328. doi:10.1242/dev.127860

Bailey, A.P., Bhattacharyya, S., Bronner-Fraser, M., Streit, A., 2006. Lens Specification Is the Ground State of All Sensory Placodes, from which FGF Promotes Olfactory Identity. Developmental Cell 11, 505–517. doi:10.1016/j.devcel.2006.08.009

Bailey, A.P., Streit, A., 2005. Sensory Organs: Making and Breaking the Pre-Placodal Region, in: Current Topics in Developmental Biology. Elsevier, pp. 167–204. doi:10.1016/S0070-2153(05)72003-2

Belecky-Adams, T.L., Adler, R., Beebe, D.C., 2002. Bone morphogenetic protein signaling and the initiation of lens fiber cell differentiation. Development 129, 3795–3802.

Bibliowicz, J., Tittle, R.K., Gross, J.M., 2011. Toward a Better Understanding of Human Eye Disease, in: Progress in Molecular Biology and Translational Science, Progress in Molecular Biology and Translational Science. Elsevier, pp. 287–330. doi:10.1016/B978-0-12-384878-9.00007-8

Cheng, S., Shakespeare, T., Mui, R., White, T.W., Valdimarsson, G., 2004. Connexin 48.5 is required for normal cardiovascular function and lens development in zebrafish embryos. J Biol Chem 279, 36993–37003. doi:10.1074/jbc.M401355200

Clemens, D.M., Németh-Cahalan, K.L., Trinh, L., Zhang, T., Schilling, T.F., Hall, J.E., 2013. In vivo analysis of aquaporin 0 function in zebrafish: permeability regulation is required for lens transparency. Invest Ophthalmol Vis Sci 54, 5136–5143. doi:10.1167/iovs.13-12337

Coolen, M., Sii-Felice, K., Bronchain, O., Mazabraud, A., Bourrat, F., Rétaux, S., Felder-Schmittbuhl, M.P., Mazan, S., Plouhinec, J.L., 2005. Phylogenomic analysis and expression patterns of large Maf genes in Xenopus tropicalis provide new insights into the functional evolution of the gene family in osteichthyans. Dev Genes Evol 215, 327–339. doi:10.1007/s00427-005-0476-y

Cui, W., Tomarev, S.I., Piatigorsky, J., Chepelinsky, A.B., Duncan, M.K., 2004. Mafs, Prox1, and Pax6 can regulate chicken betaB1-crystallin gene expression. J Biol Chem 279, 11088–11095. doi:10.1074/jbc.M312414200

Cvekl, A., Duncan, M.K., 2007. Genetic and epigenetic mechanisms of gene regulation during lens development. Progress in retinal and eye research 26, 555–597. doi:10.1016/j.preteyeres.2007.07.002

Cvekl, A., Zhang, Xin, 2017. Signaling and Gene Regulatory Networks in Mammalian Lens

Development. Trends Genet 33, 677–702. doi:10.1016/j.tig.2017.08.001
Dash, S., Brastrom, L.K., Patel, S.D., Scott, C.A., Slusarski, D.C., Lachke, S.A., 2020. The master transcription factor SOX2, mutated in anophthalmia/microphthalmia, is post-transcriptionally regulated by the conserved RNA-binding protein RBM24 in vertebrate eye development. Hum Mol Genet 29, 591–604. doi:10.1093/hmg/ddz278
Dash, S., Dang, C.A., Beebe, D.C., Lachke, S.A., 2015. Deficiency of the RNA binding protein caprin2 causes lens defects and features of peters anomaly. Dev Dyn 244, 1313–1327. doi:10.1002/dvdy.24303
Donner, A.L., Ko, F., Episkopou, V., Maas, R.L., 2007. Pax6 is misexpressed in Sox1 null lens fiber cells. Gene Expression Patterns 7, 606–613. doi:10.1016/j.modgep.2007.01.001
Eden, E., Lipson, D., Yogev, S., Yakhini, Z., 2007. Discovering Motifs in Ranked Lists of DNA Sequences. PLoS Comp Biol 3, e39–15. doi:10.1371/journal.pcbi.0030039
Eden, E., Navon, R., Steinfeld, I., Lipson, D., Yakhini, Z., 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 25–7. doi:10.1186/1471-2105-10-48
Elicker, K.S., Hutson, L.D., 2007. Genome-wide analysis and expression profiling of the small heat shock proteins in zebrafish. Gene 403, 60–69. doi:10.1016/j.gene.2007.08.003
Farnsworth, D.R., Saunders, L., Miller, A.C., 2019. A Single-Cell Transcriptome Atlas for Zebrafish Development 1–33. doi:10.1101/738344
FitzGerald, P.G., 1990. Methods for the Circumvention of Problems Associated with the Study of the Ocular Lens Plasma Membrane-Cytoskeleton Complex. Current Eye Research 9, 1083–1097. doi:10.3109/02713689008997582
FitzGerald, P.G., Gottlieb, W., 1989. The M r115 kd fiber cell-specific protein is a component of the lens cytoskeleton. Current Eye Research 8, 801–811. doi:10.3109/02713688909000870
French, C.R., Erickson, T., Callander, D., Berry, K.M., Koss, R., Hagey, D.W., Stout, J., Wuennenberg-Stapleton, K., Ngai, J., Moens, C.B., Waskiewicz, A.J., 2007. Pbx homeodomain proteins pattern both the zebrafish retina and tectum. BMC Dev Biol 7, 85–12. doi:10.1186/1471-213X-7-85
Froger, A., Clemens, D., Kalman, K., Németh-Cahalan, K.L., Schilling, T.F., Hall, J.E., 2010. Two distinct aquaporin 0s required for development and transparency of the zebrafish lens. Invest Ophthalmol Vis Sci 51, 6582–6592. doi:10.1167/iovs.10-5626
Gao, M., Huang, Y., Wang, L., Huang, M., Liu, F., Liao, S., Yu, S., Lu, Z., Han, S., Hu, X., Qu, Z., Liu, X., Yimer, T.A., Yang, L., Tang, Z., Li, D.W.-C., Liu, M., 2017. HSF4 regulates lens fiber cell differentiation by activating p53 and its downstream regulators. Cell Death Dis 1–12. doi:10.1038/cddis.2017.478
Gerber, V., Yang, L., Takamiya, M., Ribes, V., Gourain, V., Peravali, R., Stegmaier, J., Mikut, R., Reischl, M., Ferg, M., Rastegar, S., Strähle, U., 2019. The HMG box transcription factors Sox1a and Sox1b specify a new class of glycinergic interneuron in the spinal cord of zebrafish embryos. Development 146, dev172510–13. doi:10.1242/dev.172510
Giovannone, A.J., Winterstein, C., Bhattaram, P., Reales, E., Low, S.H., Baggs, J.E., Xu, M., Lalli, M.A., Hogenesch, J.B., Weimbs, T., 2018. Soluble syntaxin 3 functions as a transcriptional regulator. Journal of Biological Chemistry 293, 5478–5491. doi:10.1074/jbc.RA117.000874
Gongal, P.A., March, L.D., Holly, V.L., Pillay, L.M., Berry-Wynne, K.M., Kagechika, H., Waskiewicz, A.J., 2011. Hmx4 regulates Sonic hedgehog signaling through control of retinoic acid synthesis during forebrain patterning. Dev Biol 355, 55–64. doi:10.1016/j.ydbio.2011.04.018
Gounari, F., Merdes, A., Quinlan, R., Hess, J., FitzGerald, P.G., Ouzounis, C.A., Georgatos, S.D., 1993. Bovine filensin possesses primary and secondary structure similarity to intermediate

filament proteins. J Cell Biol 121, 847–853. doi:10.1083/jcb.121.4.847
Grainger, R.M., 1992. Embryonic lens induction: shedding light on vertebrate tissue determination. Trends in Genetics 8, 349–355. doi:10.1016/0168-9525(92)90280-H
Greiling, T.M.S., Aose, M., Clark, J.I., 2010. Cell fate and differentiation of the developing ocular lens. Invest Ophthalmol Vis Sci 51, 1540–1546. doi:10.1167/iovs.09-4388
Greiling, T.M.S., Clark, J.I., 2009. Early lens development in the zebrafish: a three-dimensional time-lapse analysis. Dev Dyn 238, 2254–2265. doi:10.1002/dvdy.21997
Greiling, T.M.S., Houck, S.A., Clark, J.I., 2009. The zebrafish lens proteome during development and aging. Mol Vis 15, 2313–2325.
Grifone, R., Saquet, A., Xu, Z., Shi, D.-L., 2018. Expression patterns of Rbm24 in lens, nasal epithelium, and inner ear during mouse embryonic development. Dev Dyn 247, 1160–1169. doi:10.1002/dvdy.24666
Grocott, T., Tambalo, M., Streit, A., 2012. The peripheral sensory nervous system in the vertebrate head A gene regulatory perspective. Dev Biol 370, 3–23. doi:10.1016/j.ydbio.2012.06.028
Hall, J.E., Mathias, R.T., 2014. The Aquaporin Zero Puzzle. Biophys J 107, 10–15. doi:10.1016/j.bpj.2014.05.023
Hess, J.F., Casselman, J.T., FitzGerald, P.G., 1996. Gene structure and cDNA sequence identify the beaded filament protein CP49 as a highly divergent type I intermediate filament protein. J Biol Chem 271, 6729–6735. doi:10.1074/jbc.271.12.6729
Horwitz, J., 1992. Alpha-crystallin can function as a molecular chaperone. Proc Natl Acad Sci USA 89, 10449–10453.
Kakrana, A., Yang, A., Anand, D., Djordjevic, D., Ramachandruni, D., Singh, A., Huang, H., Ho, J.W.K., Lachke, S.A., 2017. iSyTE 2.0: a database for expression-based gene discovery in the eye. Nucleic Acids Res 46, D875–D885. doi:10.1093/nar/gkx837
Kim, J.I., Li, T., Ho, I.C., Grusby, M.J., Glimcher, L.H., 1999. Requirement for the c-Maf transcription factor in crystallin gene regulation and lens development. Proc Natl Acad Sci USA 96, 3781–3785. doi:10.1073/pnas.96.7.3781
Kiss, A.J., Cheng, C.-H.C., 2008. Molecular diversity and genomic organisation of the alpha, beta and gamma eye lens crystallins from the Antarctic toothfish Dissostichus mawsoni. Comp Biochem Physiol Part D Genomics Proteomics 3, 155–171. doi:10.1016/j.cbd.2008.02.002
Krall, M., Lachke, S.A., 2018. A zebrafish model of foxe3 deficiency demonstrates lens and eye defects with dysregulation of key genes involved in cataract formation in humans. Hum. Genet. 137, 315–328. doi:10.1007/s00439-018-1884-1
Lachke, S.A., Alkuraya, F.S., Kneeland, S.C., Ohn, T., Aboukhalil, A., Howell, G.R., Saadi, I., Cavallesco, R., Yue, Y., Tsai, A.C.-H., Nair, K.S., Cosma, M.I., Smith, R.S., Hodges, E., AlFadhli, S.M., Al-Hajeri, A., Shamseldin, H.E., Behbehani, A., Hannon, G.J., Bulyk, M.L., Drack, A.V., Anderson, P.J., John, S.W.M., Maas, R.L., 2011. Mutations in the RNA Granule Component TDRD7 Cause Cataract and Glaucoma. Science 331, 1571–1576. doi:10.1126/science.1195970
Lampi, K.J., Wilmarth, P.A., Murray, M.R., David, L.L., 2014. Lens β-crystallins: the role of deamidation and related modifications in aging and cataract. Prog Biophys Mol Biol 115, 21–31. doi:10.1016/j.pbiomolbio.2014.02.004
Landgren, H., Blixt, A., Carlsson, P., 2008. Persistent FoxE3 expression blocks cytoskeletal remodeling and organelle degradation during lens fiber differentiation. Invest Ophthalmol Vis Sci. doi:10.1167/iovs.08-2243
Lee, H.-C., Fu, C.-Y., Zeng, C.-W., Tsai, H.-J., 2017. Embryonic expression patterns of Eukaryotic

EndoU ribonuclease family gene endouC in zebrafish. Gene Expression Patterns 25-26, 66–70. doi:10.1016/j.gep.2017.06.003
Limi, S., Zhao, Y., Guo, P., Lopez-Jones, M., Zheng, D., Singer, R.H., Skoultchi, A.I., Cvekl, A., 2019. Bidirectional Analysis of Cryba4-Crybb1 Nascent Transcription and Nuclear Accumulation of Crybb3 mRNAs in Lens Fibers. Invest Ophthalmol Vis Sci 60, 234–244. doi:10.1167/iovs.18-25921
Liu, D., Yin, B., Wang, Q., Ju, W., Chen, Y., Qiu, H., Li, J., Peng, X., Lu, C., 2012. Cytoplasmic poly(A) binding protein 4 is highly expressed in human colorectal cancer and correlates with better prognosis. J Genet Genomics 39, 369–374. doi:10.1016/j.jgg.2012.05.007
Liu, J., Zhou, Y., Qi, X., Chen, J., Chen, W., Qiu, G., Wu, Z., Wu, N., 2016. CRISPR/Cas9 in zebrafish: an efficient combination for human genetic diseases modeling. Hum. Genet. 136, 1–12. doi:10.1007/s00439-016-1739-6
Mao, L., Shelden, E.A., 2006. Developmentally regulated gene expression of the small heat shock protein Hsp27 in zebrafish embryos. Gene Expr Patterns 6, 127–133. doi:10.1016/j.modgep.2005.07.002
Mishra, S., Wu, S.-Y., Fuller, A.W., Wang, Z., Rose, K.L., Schey, K.L., McHaourab, H.S., 2018. Loss of αB-crystallin function in zebrafish reveals critical roles in the development of the lens and stress resistance of the heart. J Biol Chem 293, 740–753. doi:10.1074/jbc.M117.808634
Morris, A.C., 2011. The genetics of ocular disorders: Insights from the zebrafish. Birth Defects Res. C Embryo Today 93, 215–228. doi:10.1002/bdrc.20211
Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R., Episkopou, V., 1998. Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice. Genes Dev. 12, 776–781. doi:10.1101/gad.12.6.776
Nishijima, I., Ohtoshi, A., 2006. Characterization of a novel *prospero*-related homeobox gene, *Prox2*. Mol Genet Genomics 275, 471–478. doi:10.1007/s00438-006-0105-0
Ogden, A.T., Nunes, I., Ko, K., Wu, S., Hines, C.S., Wang, A.F., Hegde, R.S., Lang, R.A., 1998. GRIFIN, a novel lens-specific protein related to the galectin family. J Biol Chem 273, 28889–28896.
Pistocchi, A., Bartesaghi, S., Cotelli, F., Del Giacco, L., 2008. Identification and expression pattern of zebrafish prox2during embryonic development. Dev Dyn 237, 3916–3920. doi:10.1002/dvdy.21798
Pistocchi, A., Feijóo, C.G., Cabrera, P., Villablanca, E.J., Allende, M.L., Cotelli, F., 2009. The zebrafish prospero homolog prox1 is required for mechanosensory hair cell differentiation and functionality in the lateral line. BMC Dev Biol 9, 58–10. doi:10.1186/1471-213X-9-58
Posner, M., Hawke, M., Lacava, C., Prince, C.J., Bellanco, N.R., Corbin, R.W., 2008. A proteome map of the zebrafish (Danio rerio) lens reveals similarities between zebrafish and mammalian crystallin expression. Mol Vis 14, 806–814.
Posner, M., Kantorow, M., Horwitz, J., 1999. Cloning, sequencing and differential expression of alphaB-crystallin in the zebrafish, Danio rerio. Biochim Biophys Acta 1447, 271–277.
Posner, M., Murray, K.L., McDonald, M.S., Eighinger, H., Andrew, B., Drossman, A., Haley, Z., Nussbaum, J., David, L.L., Lampi, K.J., 2017. The zebrafish as a model system for analyzing mammalian and native α-crystallin promoter function. PeerJ 5, e4093–27. doi:10.7717/peerj.4093
Posner, M., Skiba, J., Brown, M., Liang, J.O., Nussbaum, J., Prior, H., 2013. Loss of the small heat shock protein αA-crystallin does not lead to detectable defects in early zebrafish lens development. Exp Eye Res. doi:10.1016/j.exer.2013.09.007
R Core Team, n.d. R: A Language and Environment for Statistical Computing.

Reza, H.M., Yasuda, K., 2004. Roles of Maf family proteins in lens development. Dev Dyn 229, 440–448. doi:10.1002/dvdy.10467
Richardson, R., Tracey-White, D., Webster, A., Moosajee, M., 2016. The zebrafish eye—a paradigm for investigating human ocular genetics. Nature Publishing Group 1–19. doi:10.1038/eye.2016.198
Robinson, M.L., Overbeek, P.A., 1996. Differential expression of alpha A- and alpha B-crystallin during murine ocular development. Invest Ophthalmol Vis Sci 37, 2276–2284.
Runkle, S., Hill, J., Kantorow, M., Horwitz, J., Posner, M., 2002. Sequence and spatial expression of zebrafish (Danio rerio) alphaA-crystallin. Mol Vis 8, 45–50.
Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., Regev, A., 2015. Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. 33, 495–502. doi:10.1038/nbt.3192
Shao, M., Lu, T., Zhang, C., Zhang, Y.-Z., Kong, S.-H., Shi, D.-L., 2020. Rbm24 controls poly(A) tail length and translation efficiency of crystallin mRNAs in the lens via cytoplasmic polyadenylation. Proceedings of the National Academy of Sciences 117, 7245–7254. doi:10.1073/pnas.1917922117
Shi, X., Bosenko, D.V., Zinkevich, N.S., Foley, S., Hyde, D.R., Semina, E.V., Vihtelic, T.S., 2005. Zebrafish pitx3 is necessary for normal lens and retinal development. Mech Dev 122, 513–527. doi:10.1016/j.mod.2004.11.012
Shi, X., Cui, B., Wang, Z., Weng, L., Xu, Z., Ma, J., Xu, G., Kong, X., Hu, L., 2009. Removal of Hsf4 leads to cataract development in mice through down-regulation of gamma S-crystallin and Bfsp expression. BMC Mol Biol 10, 10. doi:10.1186/1471-2199-10-10
Shi, X., Luo, Y., Howley, S., Dzialo, A., Foley, S., Hyde, D.R., Vihtelic, T.S., 2006. Zebrafish foxe3: roles in ocular lens morphogenesis through interaction with pitx3. Mech Dev 123, 761–782. doi:10.1016/j.mod.2006.07.004
Siddam, A.D., Gautier-Courteille, C., Perez-Campos, L., Anand, D., Kakrana, A., Dang, C.A., Legagneux, V., Méreau, A., Viet, J., Gross, J.M., Paillard, L., Lachke, S.A., 2018. The RNA-binding protein Celf1 post-transcriptionally regulates p27Kip1 and Dnase2b to control fiber cell nuclear degradation in lens development. PLoS Genet 14, e1007278–28. doi:10.1371/journal.pgen.1007278
Sinha, D., Klise, A., Sergeev, Y., Hose, S., Bhutto, I.A., Hackler, L., Malpic-Llanos, T., Samtani, S., Grebe, R., Goldberg, M.F., Hejtmancik, J.F., Nath, A., Zack, D.J., Fariss, R.N., McLeod, D.S., Sundin, O., Broman, K.W., Lutty, G.A., Zigler, J.S., 2008. betaA3/A1-crystallin in astroglial cells regulates retinal vascular remodeling during development. Mol Cell Neurosci 37, 85–95. doi:10.1016/j.mcn.2007.08.016
Sinha, D., Valapala, M., Bhutto, I., Patek, B., Zhang, C., Hose, S., Yang, F., Cano, M., Stark, W.J., Lutty, G.A., Zigler, J.S., Wawrousek, E.F., 2012. βA3/A1-crystallin is required for proper astrocyte template formation and vascular remodeling in the retina. Transgenic Res. doi:10.1007/s11248-012-9608-0
Smith, A.A., Wyatt, K., Vacha, J., Vihtelic, T.S., Zigler, J.S., Wistow, G.J., Posner, M., 2006. Gene duplication and separation of functions in alphaB-crystallin from zebrafish (Danio rerio). FEBS J 273, 481–490. doi:10.1111/j.1742-4658.2005.05080.x
Song, P., Perkins, B.D., 2018. Developmental expression of the zebrafish Arf-like small GTPase paralogs arl13a and arl13b. Gene Expression Patterns 29, 82–87. doi:10.1016/j.gep.2018.07.002
Srivastava, R., Budak, G., Dash, S., Lachke, S.A., Janga, S.C., 2017. Transcriptome analysis of developing lens reveals abundance of novel transcripts and extensive splicing alterations.

Nature Publishing Group 1–17. doi:10.1038/s41598-017-10615-4
Swan, C.L., Evans, T.G., Sylvain, N., Krone, P.H., 2012. Zebrafish HSF4: a novel protein that shares features of both HSF1 and HSF4 of mammals. Cell Stress Chaperones 17, 623–637. doi:10.1007/s12192-012-0337-3
Takeuchi, T., Kudo, T., Ogata, K., Hamada, M., Nakamura, M., Kito, K., Abe, Y., Ueda, N., Yamamoto, M., Engel, J.D., Takahashi, S., 2009. Neither MafA/L-Maf nor MafB is essential for lens development in mice. Genes Cells 14, 941–947. doi:10.1111/j.1365-2443.2009.01321.x
Tittle, R.K., Sze, R., Ng, A., Nuckels, R.J., Swartz, M.E., Anderson, R.M., Bosch, J., Stainier, D.Y.R., Eberhart, J.K., Gross, J.M., 2011. Uhrf1 and Dnmt1 are required for development and maintenance of the zebrafish lens. Dev Biol 350, 50–63. doi:10.1016/j.ydbio.2010.11.009
Tsonis, P.A., 2007. Regeneration via transdifferentiation: The lens and hair cells. Hearing Research 227, 28–31. doi:10.1016/j.heares.2006.06.011
Tsonis, P.A., Call, M.K., Grogg, M.W., Sartor, M.A., Taylor, R.R., Forge, A., Fyffe, R., Goldenberg, R., Cowper-Sal lari, R., Tomlinson, C.R., 2007. MicroRNAs and regeneration: Let-7 members as potential regulators of dedifferentiation in lens and inner ear hair cell regeneration of the adult newt. Biochemical and Biophysical Research Communications 362, 940–945. doi:10.1016/j.bbrc.2007.08.077
Vinayak, J., Marrella, S.A., Hussain, R.H., Rozenfeld, L., Solomon, K., Bayfield, M.A., 2018. Human La binds mRNAs through contacts to the poly(A) tail. Nucleic Acids Res 46, 4228–4240. doi:10.1093/nar/gky090
Vorontsova, I., Gehring, I., Hall, J.E., Schilling, T.F., 2018. Aqp0a Regulates Suture Stability in the Zebrafish Lens. Invest Ophthalmol Vis Sci 59, 2869–11. doi:10.1167/iovs.18-24044
Wages, P., Horwitz, J., Ding, L., Corbin, R.W., Posner, M., 2013. Changes in zebrafish (Danio rerio) lens crystallin content during development. Mol Vis 19, 408–417.
Westerfield, M., 2007. The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (*Danio rerio*). University of Oregon Press, Eugene.
Wigle, J.T., Chowdhury, K., Gruss, P., Oliver, G., 1999. Prox1 function is crucial for mouse lens-fibre elongation. Nat Genet 21, 318–322. doi:10.1038/6844
Wilson, C.A., High, S.K., McCluskey, B.M., Amores, A., Yan, Y.-L., Titus, T.A., Anderson, J.L., Batzel, P., Carvan, M.J., Schartl, M., Postlethwait, J.H., 2014. Wild sex in zebrafish: loss of the natural sex determinant in domesticated strains. Genetics 198, 1291–1308. doi:10.1534/genetics.114.169284
Wistow, G., 2012. The human crystallin gene families. Hum. Genomics 6, 26–10. doi:10.1186/1479-7364-6-26
Wistow, G., Wyatt, K., David, L., Gao, C., Bateman, O., Bernstein, S., Tomarev, S., Segovia, L., Slingsby, C., Vihtelic, T., 2005. gammaN-crystallin and the evolution of the betagamma-crystallin superfamily in vertebrates. FEBS J 272, 2276–2291. doi:10.1111/j.1742-4658.2005.04655.x
Wistow, G.J., Piatigorsky, J., 1988. Lens Crystallins: The Evolution and Expression of Proteins for a Highly Specialized Tissue. http://dx.doi.org/10.1146/annurev.bi.57.070188.002403 57, 479–504. doi:10.1146/annurev.bi.57.070188.002403
Yang, Y., Chauhan, B.K., Cveklova, K., Cvekl, A., 2004. Transcriptional regulation of mouse alphaB- and gammaF-crystallin genes in lens: opposite promoter-specific interactions between Pax6 and large Maf transcription factors. J Mol Biol 344, 351–368. doi:10.1016/j.jmb.2004.07.102
Zhang, Xiaofang, Zhang, F., Guo, L., Wang, Y., Zhang, P., Wang, R., Zhang, N., Chen, R., 2013. Interactome Analysis Reveals that C1QBP (complement component 1, q subcomponent

binding protein) Is Associated with Cancer Cell Chemotaxis and Metastasis. Mol Cell Proteomics 12, 3199–3209. doi:10.1074/mcp.M113.029413
Zhao, Y., Wilmarth, P.A., Cheng, C., Limi, S., Fowler, V.M., Zheng, D., David, L.L., Cvekl, A., 2018a. Proteome-transcriptome analysis and proteome remodeling in mouse lens epithelium and fibers. Exp Eye Res 179, 32–46. doi:10.1016/j.exer.2018.10.011
Zhao, Y., Zheng, D., Cvekl, A., 2018b. A comprehensive spatial-temporal transcriptomic analysis of differentiating nascent mouse lens epithelial and fiber cells. Exp Eye Res 175, 56–72. doi:10.1016/j.exer.2018.06.004
Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., Gregory, M.T., Shuga, J., Montesclaros, L., Underwood, J.G., Masquelier, D.A., Nishimura, S.Y., Schnall-Levin, M., Wyatt, P.W., Hindson, C.M., Bharadwaj, R., Wong, A., Ness, K.D., Beppu, L.W., Deeg, H.J., McFarland, C., Loeb, K.R., Valente, W.J., Ericson, N.G., Stevens, E.A., Radich, J.P., Mikkelsen, T.S., Hindson, B.J., Bielas, J.H., 2019. Massively parallel digital transcriptional profiling of single cells 1–12. doi:10.1038/ncomms14049
Zigler, J.S., Zhang, C., Grebe, R., Sehrawat, G., Hackler, L., Adhya, S., Hose, S., McLeod, D.S., Bhutto, I., Barbour, W., Parthasarathy, G., Zack, D.J., Sergeev, Y., Lutty, G.A., Handa, J.T., Sinha, D., 2011. Mutation in the A3/A1-crystallin gene impairs phagosome degradation in the retinal pigmented epithelium of the rat. J Cell Sci 124, 523–531. doi:10.1242/jcs.078790
Zilinski, C.A., Shah, R., Lane, M.E., Jamrich, M., 2005. Modulation of zebrafish pitx3 expression in the primordia of the pituitary, lens, olfactory epithelium and cranial ganglia by hedgehog and nodal signaling. genesis 41, 33–40. doi:10.1002/gene.20094
Zou, P., Wu, S.-Y., Koteiche, H.A., Mishra, S., Levic, D.S., Knapik, E., Chen, W., McHaourab, H.S., 2015. A conserved role of αA-crystallin in the development of the zebrafish embryonic lens. Exp Eye Res 138, 104–113. doi:10.1016/j.exer.2015.07.001

Figure 1

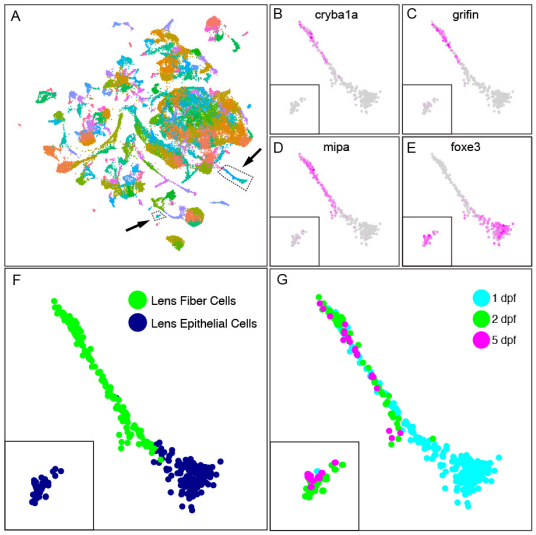

Figure 1. Expression of known lens marker genes identified lens epithelial and fiber cell clusters. Global single cell gene expression in 1, 2 and 5 dpf embryos was analyzed with UMAP to cluster cells with similar expression patterns (A). Two areas in the total atlas (A) that contained lens cells are indicated with dotted lines and arrows. Four genes known to be expressed in either lens fiber cells (*cryba1a* (B); *grifin* (C); *mipa* (D)) or epithelial cells (*foxe3* (E)) were used to identify these cell types (F). The intensity of purple coloration in panels B-E reflects gene expression. Panel G shows the developmental age of all lens cells. Note that insets in panels B-G include the 2 and 5 dpf lens epithelial cells. Gene ontology term analysis of the most expressed genes in 1 dpf compared to 2 and 5 dpf lens epithelial cells showed a shift from enrichment in protein translation, ribosome biogenesis and embryonic development at 1 dpf to membrane transport and ATP metabolism at 2 and 5 dpf (Supp Table 1).

Figure 2

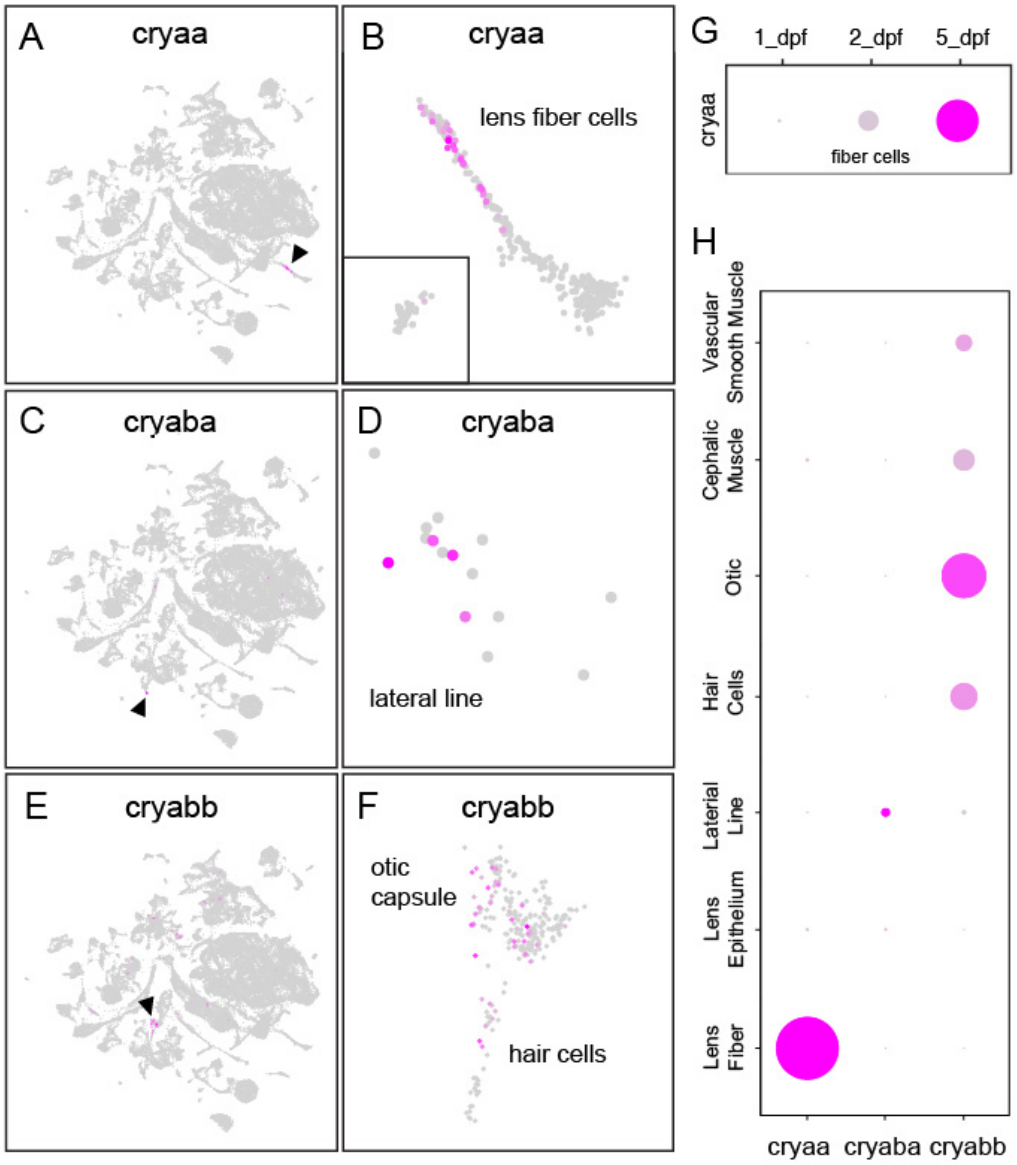

Figure 2. Lens fiber cells upregulated expression of αA-crystallin starting at 2 dpf, but αB-crystallin expression in the lens was almost non-existent through 5 dpf. mRNA for *cryaa* was found almost exclusively in the lens (A), and primarily in fiber cells (B). Expression of *cryaba* was detected primarily in lateral line cells (arrowhead in C and detailed view in D). Expression of *cryabb* was detected at relatively low levels in diverse tissues including otic capsule and hair cells (arrowhead in E and details in F). Temporal expression dynamics of cryaa expression in lens fiber cells through 5 dpf (G). Expression of α-crystallin genes in tissues with significant representation (H). For dot plots, the size of each dot reflects the percentage of cells in that tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Figure 3

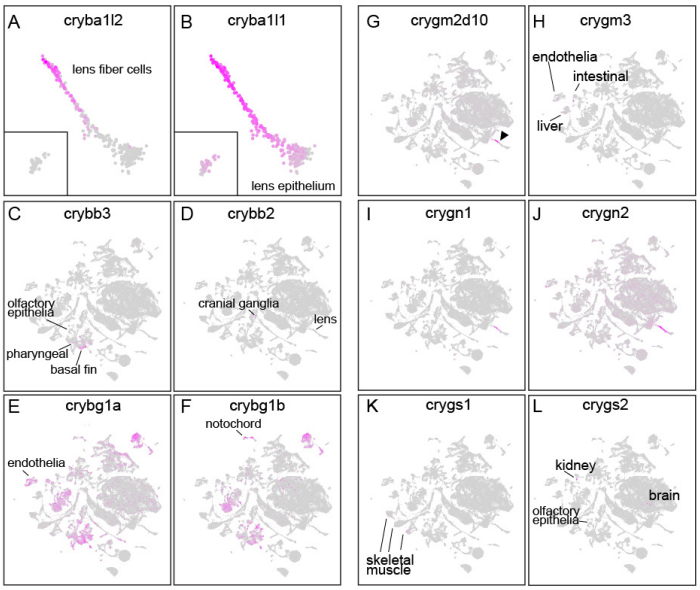

Figure 3. Spatial expression of the β-crystallin and diverse gamma crystallin gene families. Many, like *cryba1l2* were confined to lens fiber cells (A). Others like *cryba1l1* were most strongly expressed in lens fiber cells but also extended into lens epithelium (B). *crybb3* was unique among β-crystallin genes in lacking lens expression and instead being found in olfactory epithelium, pharyngeal and basal fin cells (C). *crybb2* was found in both lens fiber and cranial ganglion cells (D). The paralogs *crybg1a* and *crybg1b* had mostly overlapping expression, although the former was more abundant in endothelia (E) while the latter was abundant in notochord (F). Almost all γM-crystallin genes with detected expression were restricted to lens fiber cells, as shown for *crygm2d10* (G, arrowhead). The only γM-crystallin gene with extralenticular expression was *crygm3*, which was expressed in endothelial, intestinal epithelia and liver cells (H). Both γN-crystallin genes were expressed primarily in lens fiber cells with some lesser expression in lens epithelial cells (I and J). Both were expressed in many other tissues at weaker levels across the embryos, although this expression was stronger for *crygn2*. Only two γS-crystallin genes were expressed in our dataset. *crygs1* was expressed in a number of skeletal muscle cells (K) while *crygs2* was expressed in kidney, brain and olfactory epithelia (L). Neither γS-crystallin gene was expressed in lens.

Figure 4

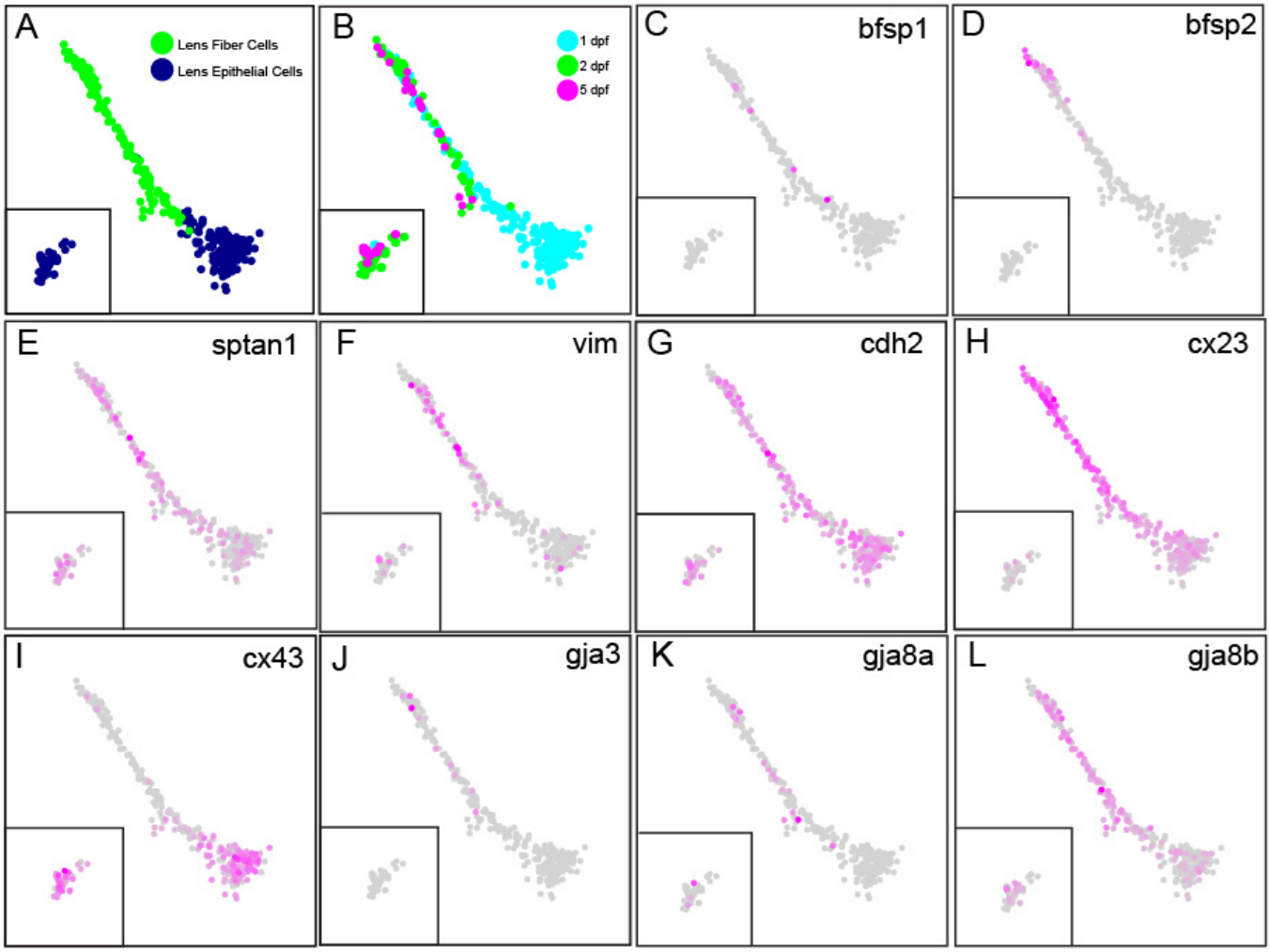

Figure 4. The zebrafish lens expresses many of the same cytoskeletal and membrane proteins as mammals, with some interesting divergence between paralogous pairs. Panels A and B indicate cell type and age as a reference (see Figure 1). Shown are the expression of two intermediate filament genes unique to the vertebrate lens, *bfsp1* (*filensin*) and *bfsp2* (*CP49/phakinin*) (A and B). Also shown is a spectrin gene (E) and the single vimentin gene (F). Multiple cytoskeletal-associated plasma membrane protein genes are also shown, including one cadherin (G) and two connexin genes (H, I). Three gap junction protein genes are shown, including *gja3* (J) and a paralog pair for *gja8a* and *gja8b* (K, L).

Figure 5

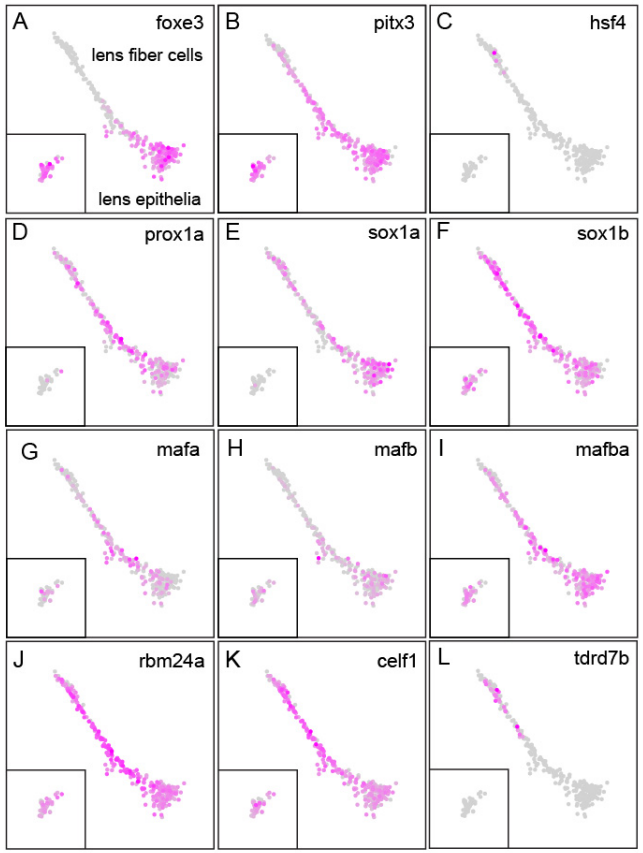

Figure 5. Spatial expression of transcription factors and RNA-binding proteins known to regulate lens development. Expression for the three transcription factors foxe3 (A), pitx3 (B) and hsf4 (C) is shown in the identified lens cell clusters. Inset contains two and five dpf epithelial cells as identified in Figure 1G. These three transcription factors have been shown to take part in shared lens development signaling pathways in both mammalian and zebrafish lens. The mammalian gene Prox1 is duplicated in zebrafish, but only the paralog prox1a is expressed in lens (D). Mammalian Sox1 also occurs as paralogs in zebrafish, and both are expressed in lens, with sox1a showing greater expression in 1 dpf epithelial cells (E) while sox1b expression appears to extend into 2 and 5 dpf epithelial cells (F: cells within inset). The expression of three large Maf family genes are shown, including the mammalian c-Maf paralogs mafa (G) and mafb (H) as well as one mammalian MafB paralog mafba (I).Three previously characterized RNA-binding protein genes known to regulate lens development were expressed in lens cells: rbm24a (J), celf1 (K) and tdrd7b (L).

Figure 6

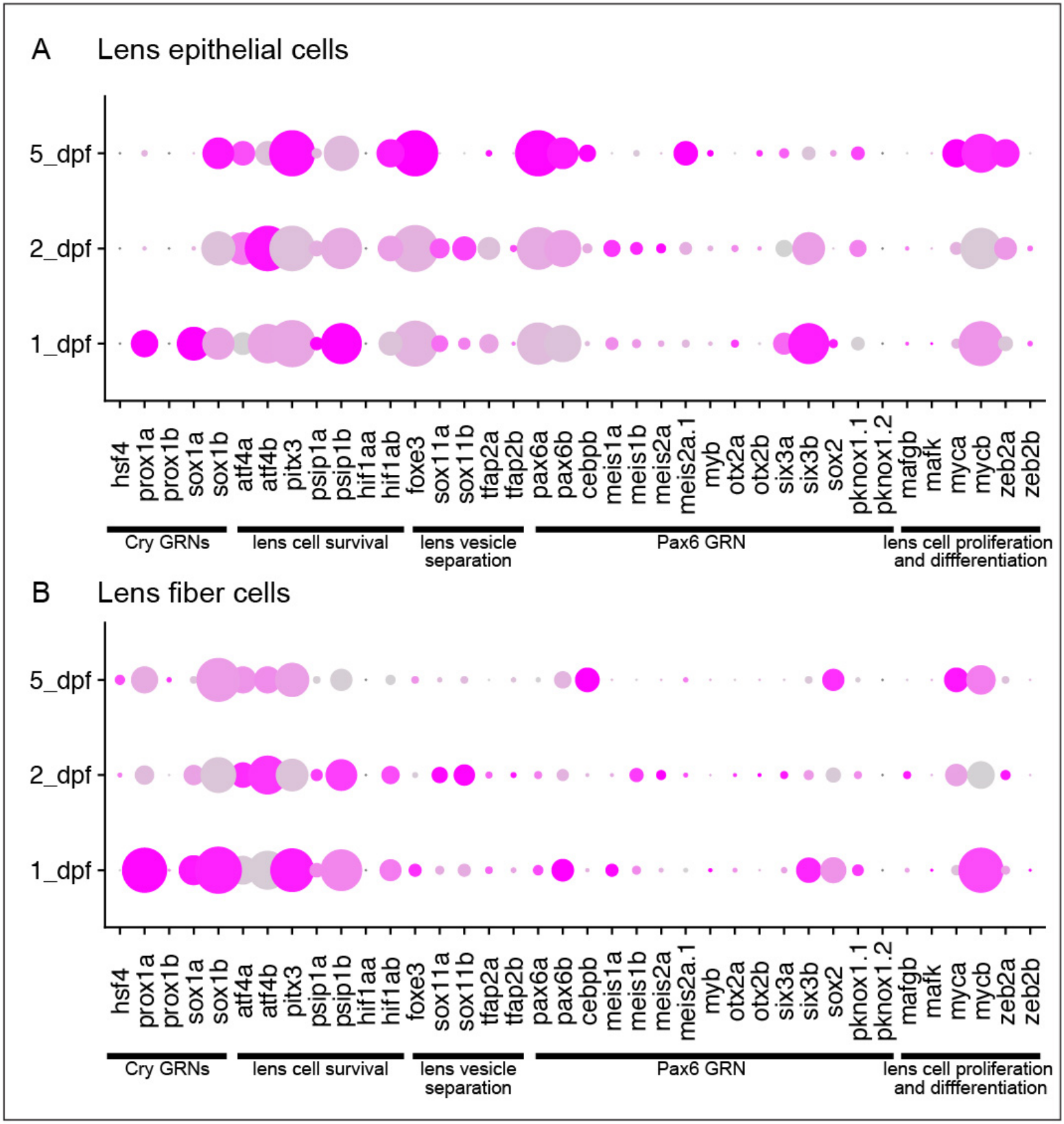

Figure 6. Temporal dynamics of transcription factor expression in zebrafish lens epithelial and fiber cells at 1, 2 and 5 dpf. Expression of lens development regulators and zebrafish paralogs for single copy mammalian genes in lens epithelial (A) and fiber cell (B) clusters. Functional categories for transcription factors are noted below each panel. The size of each dot reflects the percentage of cells in that tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Figure 7

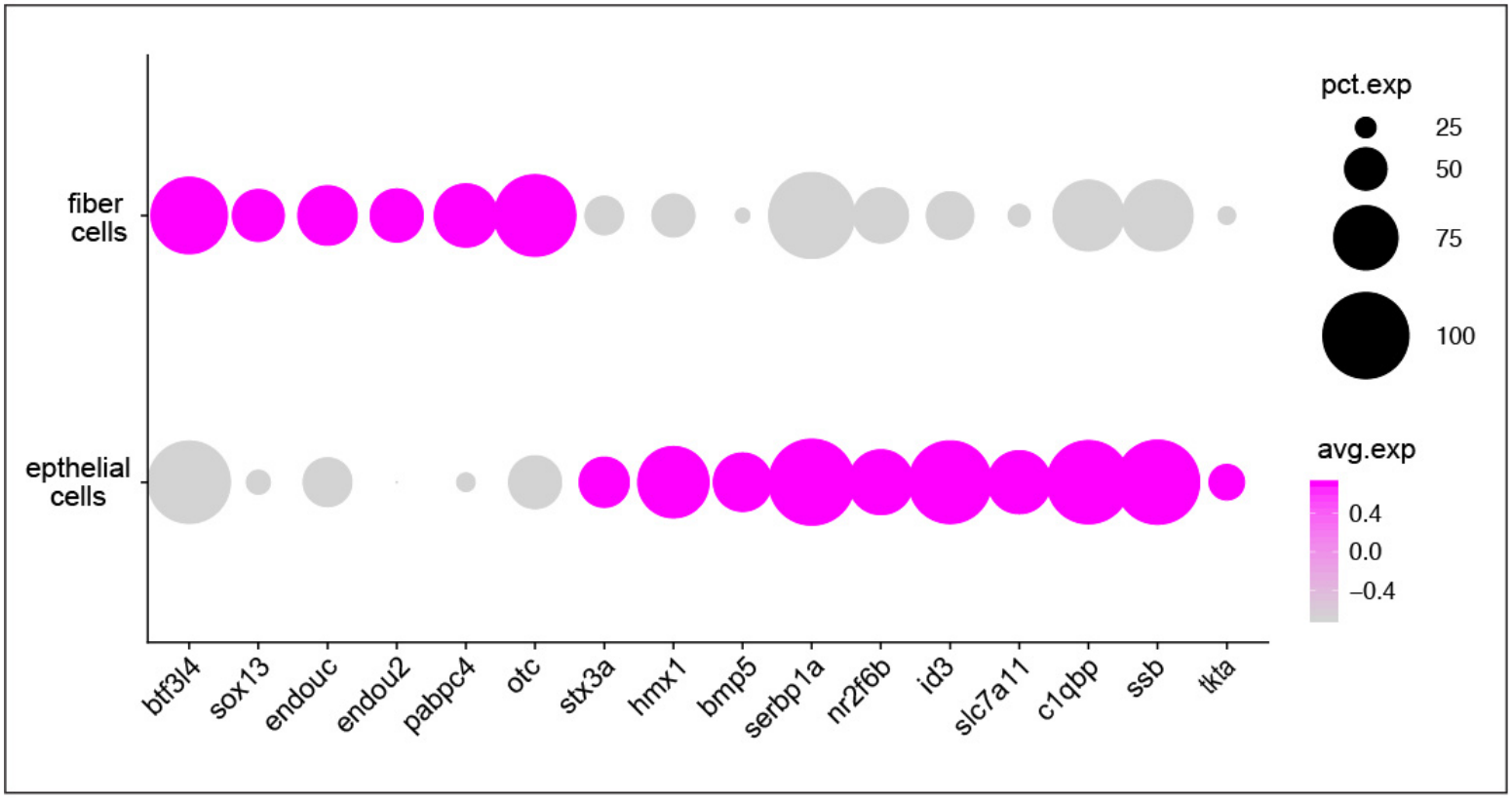

Figure 7. Expression of genes with preferred expression in lens fiber cells and epithelia, but with no currently identified lens function. The size of each dot reflects the percentage of cells in each tissue expressing the gene while intensity of magenta coloration reflects average gene expression across the tissue cluster.

Table 1. Expression of β-crystallin genes. Age of onset in hours post fertilization (hpf) determined using a publically available whole embryo RNA-Seq dataset (EBI). Location of expression is determined from our single cell RNA-Seq data. Gene expression not detected in either dataset are indicated with ND.

ZFIN Gene name	Ensemble ID	Expression onset (hpf)	Expression location
cryba1l1	00000032929	5.5	Lens fiber and epithelial cells
cryba1l2	00000014803	24	Lens fiber cells
cryba1a	00000041141	24	Lens fiber cells
cryba1b	00000053875	16	Lens fiber and epithelial cells
cryba2a	00000030349	16	Lens fiber and epithelial cells
cryba2b	00000041925	10.3	Lens fiber and epithelial cells
cryba4	00000024548	16	Lens fiber and epithelial cells
crybb1	00000100122	0	Lens fiber and epithelial cells
crybb1l1	00000007576	24	Lens fiber and epithelial cells
crybb1l2	00000016793	5.5	Lens fiber and epithelial cells
crybb1l3	00000007655	24	Lens fiber cells
crybb2	00000053512	48	Lens fiber cells; Cranial ganglia
crybb3	00000041301	16	Basal fin; Pharyngeal and olfactory epithelia
crybg1a	00000031782	5.5	Pharyngeal; Epidermal; Otic; Lateral line
crybg1b	00000006060	4.33	Pharyngeal; Epidermal; Otic; Lateral line
crybg2	00000098777	ND	ND
crybgx	00000036140	48	Lens fiber cells

Table 2. Expression of γ-crystallin genes. Age of onset in hours post fertilization (hpf) determined using a publically available whole embryo RNA-Seq dataset (EBI). Location of expression is determined from our single cell RNA-Seq data. Gene expression not detected in either dataset are indicated with ND.

ZFIN Gene name	Ensemble ID	Expression onset (hpf)	Expression location
crygm1	00000016421	ND	Skeletal muscle
crygm1b	00000040738	96	Not in Database
crygm2a	00000057536	ND	ND
crygm2b	00000115234	ND	ND
crygm2c	00000117097	ND	Lens fiber cell, iridiphore
crygm2d1	00000087324	30	Lens fiber cells
crygm2d2	00000086917	30	Lens fiber cells
crygm2d3	00000088823	30	Lens fiber cells
crygm2d4	00000087164	30	Lens fiber cells
crygm2d5	00000069792	30	Lens fiber cells
crygm2d6	00000073874	30	Lens fiber cells
crygm2d7	00000076572	30	Lens fiber cells
crygm2d8	00000116164	ND	Lens fiber cells
crygm2d9	00000115701	ND	Lens fiber cells
crygm2d10	00000087765	24	Lens fiber cells
crygm2d11	00000069827	36	Not in Database
crygm2d12	00000069801	30	Lens fiber cells
crygm2d13	00000057460	24	Lens fiber cells
crygm2d14	00000087301	30	Lens fiber cells
crygm2d15	00000069826	30	Lens fiber cells
crygm2d16	00000076790	30	Lens fiber cells
crygm2d17	00000069817	ND	Lens fiber cells
crygm2d20	00000091148	30	Lens fiber cells
crygm2d21	00000086658	30	Lens fiber cells
crygm2e	00000044875	72	Lens fiber cells
crygm2f	00000078134	72	Lens fiber cells
crygm3	00000078189	ND	Hepatic and intestinal epithelium
crygm4	00000008219	ND	ND
crygm5	00000041179	48	ND
crygm6	00000040735	ND	ND
crygm7	00000040736	ND	ND
crygmx	00000053862	24	Lens fiber and epithelial cells
crygmxl1	00000018797	36	Lens fiber cells
crygmxl2/crygmxc	00000074001	24	Lens fiber and epithelial cells
crygn1	00000102129	36	Lens fiber and epithelial cells; Hepatic and intestinal epithelium
crygn2	00000030411	4.33	Lens fiber and epithelial cells; Wide spread weak expression
crygs1	00000042992	48	Skeletal muscle

crygs2	00000041171	4.33	Olfactory, intestine, kidney
crygs3	00000025943	ND	ND
crygs4	00000010701	ND	ND

Table 3. Expression of lens cytoskeletal and membrane protein genes. Age of onset in hours post fertilization (hpf) determined using a publically available whole embryo RNA-Seq dataset (EBI). Location of expression is determined from our single cell RNA-Seq data.

ZFIN Gene name	Ensemble ID	Expression onset (hpf)	Expression location
Cytoskeletal protein genes			
bfsp1	00000074919	19	Lens fiber and epithelial cells
bfsp2	00000011998	19	Lens fiber cells
sptan1	00000019231	0	Lens fiber and epithelial cells, cranial ganglia, melanophores, xanthophores, endothelia and notochord
sptbn1	00000102883	0	Lens fiber and epithelial cells, cranial ganglia, melanophores, xanthophores, endothelia and notochord
sptbn2	00000053956	0	Lens fiber and epithelial cells, nervous tissue
sptbn5	00000097683	30	Lens fiber and epithelial cells, kidney
sptb	00000030490	16	Red Blood Cells
vim	00000010008	2.25	Lens fiber and epithelial cells, widespread
Membrane associated protein genes			
cdh2	00000018693	0	Lens fiber and epithelial cells, widespread
cx23	00000054150	0.75	Lens fiber and epithelial cells
cx43	00000041799	0	Lens fiber and epithelial cells, widespread
gja3	00000021889	19	Lens fiber cells
gja8a	00000069451	30	Lens fiber and epithelial cells
gja8b	00000015076	19	Lens fiber and epithelial cells

Table 4. Candidate novel lens-related genes. These genes were identified as having preferred expression in either lens fiber or epithelial cells, or both. The functional category of each gene is listed. Preferred expression was defined as the top 200 expressed gene in lens fiber or epithelial cells. The expression of these genes in lens fiber and epithelial cells is visualized in Figure 7.

ZFIN Gene name	Ensemble ID	Lens location	Functional Category
sox13	00000030297	Fiber	Transcription factor
id3	00000054823	Epithelia	Transcription factor
nr2f6b	00000003165	Epithelia	Transcription factor
btf3l4	00000089681	Fiber and Epithelial	Transcription factor
hmx1	00000095651	Fiber and Epithelial	Transcription factor
endou2	00000020122	Fiber	RNA-binding protein
pabpc4	00000059259	Fiber	RNA-binding protein
serbp1a	00000074242	Epithelia	RNA-binding protein
c1qbp	00000039887	Fiber and Epithelial	RNA-binding protein
endouc	00000112234	Fiber and Epithelial	RNA-binding protein
ssb	00000029252	Fiber and Epithelial	RNA-binding protein
bmp5	00000101701	Epithelia	Signal
slc7a11	00000071384	Fiber and Epithelial	Transport
stx3a	00000001880	Fiber and Epithelial	Protein transport/Transcription factor?
otc	00000062147	Fiber and Epithelial	Metabolism
tkta	00000029689	Fiber and Epithelial	Metabolism